PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :		(11) International Publication Number: WO 00/52022
C07H 19/00, 21/00, 21/02, 21/04, C12Q 1/00, G01N 33/53, C12P 21/06, 15/09, C12N 15/00, 15/09, 15/63, 15/70, 15/74, 5/00, 5/02, 1/00, 15/63, 15/85, 15/87, C07K 1/00, 14/00, 17/00	A1	(43) International Publication Date: 8 September 2000 (08.09.00)

(21) International Application Number: PCT/US00/05226

(22) International Filing Date: 1 March 2000 (01.03.00)

(30) Priority Data:
60/122,458 1 March 1999 (01.03.99) US

(71) Applicant: MILLENNIUM PHARMACEUTICALS, INC. [US/US]; 75 Sidney Street, Cambridge, MA 02139 (US).

(72) Inventors: BARNES, Thomas, M.; 22 Hanson Street #2, Boston, MA 02118 (US). HOLTZMAN, Douglas, A.; 821 Centre Street #6, Jamaica Plain, MA 02130 (US). SHARP, John, D.; 245 Park Avenue, Arlington, MA 02476-7439 (US). FRASER, Christopher, C.; 52 Grassland Street, Lexington, MA 02421 (US).

(74) Agents: CORUZZI, Laura, A. et al.; Pennie & Edmonds LLP, 1155 Avenue of the Americas, New York, NY 10036 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: SECRETED PROTEINS AND NUCLEIC ACIDS ENCODING THEM

(57) Abstract

The invention provides isolated nucleic acid molecules, designated TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267. These nucleic acid molecules encode wholly secreted and transmembrane proteins. The invention also provides antisense nucleic acid molecules, expression vectors containing the nucleic acid molecules of the invention, host cells into which the expression vectors have been introduced, and nonhuman transgenic animals in which a nucleic acid molecule of the invention has been introduced or disrupted. The invention still further provides isolated polypeptides, fusion polypeptides, antigenic peptides and antibodies. Diagnostic, screening and therapeutic methods utilizing compositions of the invention are also provided.

```
                   CTCCACTCCTAGTTTGTTGTGCCGTCGCCGCGTCCCAGGGACCCCTTGTCCCGAAGCGCACG   62
GCAGCGGGGCGGACTTCAGCCCTCCAGCCGGGCTGCGTTCCAGGTCCCGGTCCGAGGCGGCGCTGGAGCCTCGGCCCC  141
AGGCCCGGACAGGAACTCCTTTCGCGAGCCTGCCCCAGCGCGCGGCGTCTCAGCTGCTCGCCGCCCCCCACCCCAGAGTGC  220
GTGCCGGGTGACTCCCGCCACCTTTGCCGACCCTCCTGACCTTAGCGGACTGCCAGCCCGAGGGAGTCTCAGGCCCCCGC  299

         M   V   A   E   R   S   P   A   R   S   P   G   S   W   L   F   P   G     18
CCGCAGG ATG GTG GCG GAG CGG TCC CCG GCC CGC AGC CCC GGG AGC TGG CTG TTC CCC GGG   360

 L   W   L   L   V   L   S   G   P   G   G   L   L   R   A   Q   E   Q   P   S     38
CTG TGG CTG TTG GTG CTC AGC GGT CCC GGG GGG CTG CTG CGC GCC CAG GAG CAG CCC TCC   420

 C   R   R   A   F   D   L   Y   F   V   L   D   K   S   G   S   V   A   N   N     58
TGC AGA AGA GCC TTT GAT CTC TAC TTC GTC CTG GAC AAG TCT GGG AGT GTG GCA AAT AAC   480

 W   I   E   I   Y   N   F   V   Q   Q   L   A   E   R   F   V   S   P   E   M     78
TGG ATT GAA ATT TAT AAT TTC GTA CAG CAA CTT GCG GAG AGA TTT GTG AGC CCT GAA ATG   540

 R   L   S   F   I   V   F   S   S   Q   A   T   I   I   L   P   L   T   G   D     98
AGA TTA TCT TTC ATT GTG TTT TCT TCT CAA GCA ACT ATT ATT TTG CCA TTA ACT GGA GAC   600

 R   G   K   I   S   K   G   L   E   D   L   K   R   V   S   P   V   G   E   T    118
AGA GGC AAA ATC AGT AAA GGC TTG GAG GAT TTA AAA CGT GTT AGT CCA GTA GGA GAG ACA   660

 Y   I   H   E   G   L   K   L   A   N   E   Q   I   Q   K   A   G   G   L   K    138
TAT ATC CAT GAA GGA CTA AAG CTA GCG AAT GAA CAA ATT CAG AAA GCA GGA GGC TTG AAA   720

 T   S   S   I   I   I   A   L   T   D   G   K   L   D   G   L   V   P   S   Y    158
ACC TCC AGT ATC ATA ATT GCT CTG ACA GAT GGC AAG TTG GAC GGT CTG GTG CCA TCA TAT   780

 A   E   K   E   A   K   I   S   R   S   L   G   A   S   V   Y   C   V   G   V    178
GCA GAG AAA GAG GCA AAG ATA TCC AGG TCA CTT GGG GCT AGT GTT TAT TGT GTT GGT GTC   840

 L   D   F   E   Q   A   Q   L   E   R   I   A   D   S   K   E   Q   V   F   P    198
CTT GAT TTT GAA CAA GCA CAG CTT GAA AGA ATT GCT GAT TCC AAG GAG CAA GTT TTC CCT   900

 V   K   G   G   F   Q   A   L   K   G   I   I   N   S   I   L   A   Q   S   C    218
GTC AAA GGT GGA TTT CAG GCT CTT AAA GGA ATA ATT AAT TCT ATA CTA GCT CAG TCA TGT   960

 T   E   I   L   E   L   Q   P   S   S   V   C   V   G   E   E   F   Q   I   V    238
ACT GAA ATC CTA GAA TTG CAG CCC TCA AGT GTC TGT GTG GGG GAG GAA TTT CAG ATT GTC  1020

 L   S   G   R   G   F   M   L   G   S   R   N   G   S   V   L   C   T   Y   T    258
TTA AGT GGA AGA GGA TTC ATG CTG GGC AGT CGG AAT GGC AGT GTT CTC TGC ACT TAC ACT  1080
```

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

SECRETED PROTEINS AND NUCLEIC ACIDS ENCODING THEM

Cross Reference to Related Applications

This is a continuation-in-part application which claims the benefit of priority under 35 U.S.C. § 119(e) of co-pending provisional United States Patent Application No. 60/122,458, filed March 1, 1999, the entire contents of which is incorporated herein by reference in its entirety.

Background of the Invention

Many secreted proteins, for example, cytokines and cytokine receptors, play a vital role in the regulation of cell growth, cell differentiation, and a variety of specific cellular responses. A number of medically useful proteins, including erythropoietin, granulocyte-macrophage colony stimulating factor, human growth hormone, and various interleukins, are secreted proteins. Thus, an important goal in the design and development of new therapies is the identification and characterization of secreted and transmembrane proteins and the genes which encode them.

Many secreted proteins are receptors which bind a ligand and transduce an intracellular signal, leading to a variety of cellular responses. The identification and characterization of such a receptor enables one to identify both the ligands which bind to the receptor and the intracellular molecules and signal transduction pathways associated with the receptor, permitting one to identify or design modulators of receptor activity, e.g., receptor agonists or antagonists and modulators of signal transduction.

Summary of the Invention

The present invention is based, at least in part, on the discovery of cDNA molecules encoding TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267, all of which are either wholly secreted or transmembrane proteins. These proteins, fragments, derivatives, and variants thereof are collectively referred to as polypeptides of the invention or proteins of the invention. Nucleic acid molecules encoding polypeptides of the invention are collectively referred to as nucleic acids of the invention.

The nucleic acids and polypeptides of the present invention are useful as modulating agents in regulating a variety of cellular processes. Accordingly, in one aspect, the present invention provides isolated nucleic acid molecules encoding a polypeptide of the invention or a biologically active portion thereof. The present invention also provides nucleic acid

molecules which are suitable as primers or hybridization probes for the detection of nucleic acids encoding a polypeptide of the invention.

The invention features nucleic acid molecules which are at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the nucleotide sequence of any of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the nucleotide sequence of the cDNA of a clone deposited with the American Type Culture Collection (ATCC™) as Accession Number 207176, or a complement thereof.

The invention features nucleic acid molecules which are at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the nucleotide sequence of any of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the nucleotide sequence of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof, wherein such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention.

The invention features nucleic acid molecules of at least 625, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500, 3600, or 3677 nucleotides of the nucleotide sequence of SEQ ID NO:1, the nucleotide sequence of the human TANGO 216 cDNA clone of ATCC™ Accession No. 207176, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 600, 650, 700, 750, 800, 850, 900, 950, 1000, or 1040 nucleotides of nucleic acids 1695 to 2737 of SEQ ID NO:1, or a complement thereof, wherein such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention.

The invention features nucleic acid molecules which include a fragment of at least 625, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450, or 1460 nucleotides of the nucleotide sequence of SEQ ID NO:3, or a complement thereof.

The invention features nucleic acid molecules of at least 675, 700, 725, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2300, 2400, 2500, 2600, 2700, 2800, 2900, 3000, 3100, 3200, 3300, 3400, 3500 or 3501 nucleotides of the nucleotide sequence of SEQ ID NO:16, the nucleotide sequence of a mouse TANGO 216 cDNA, or a complement thereof. The invention features nucleic acid molecules comprising at least 85, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350,

375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1775, 1800, 1825, 1850, 1875, 1900, 1925, 1950, 1975, 2000, 2025, 2050, 2075, 2100, 2125, 2150, 2175, 2200, 2225, 2250, 2275, 2300, 2325, 2350, 2375, 2400 nucleotides of nucleic acids 1 to 2417 of SEQ ID NO:16, or a complement thereof.

The invention features nucleic acid molecules comprising at least 85, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, 1450 or 1460 nucleotides of the nucleotide sequence of SEQ ID NO:18, or a complement thereof.

The invention features nucleic acid molecules of at least 525, 550, 600, 650, 700, 750, 800, 850, 900, 950,or 969 nucleotides of the nucleotide sequence of SEQ ID NO:4, the nucleotide sequence of a human TANGO 261 cDNA, the nucleotide sequence of the human TANGO 261 cDNA clone of ATCC™ Accession No. 207176, or a complement thereof. The invention also features nucleic acid molecules comprising at least 280, 300, 320, 340, 360, 380, 400, 420, 440, 450 nucleotides of nucleic acids 1 to 453 of SEQ ID NO:4, or a complement thereof.

The invention features nucleic acid molecules of at least 280, 300, 320, 340, 360, 380, 400, 420, 440, 460, 500, 520, 540, 560, 580, 600, 620, 640, 660, 680, 700, 720, 740, or 750 nucleotides of the nucleotide sequence of SEQ ID NO:6, or a complement thereof.

The invention features nucleic acid molecules of at least 560, 575, 600, 625, 650, 675, 700, 725, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, or 1713 nucleotides of the nucleotide sequence of SEQ ID NO:19, the nucleotide sequence of a mouse TANGO 261 cDNA, or a complement thereof. The invention features nucleic acid molecules comprising at least 25 or 30 nucleotides of nucleic acids 1 to 33 of SEQ ID NO:19, or a complement thereof. The invention features nucleic acid molecules comprising at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, or 170 nucleotides of nucleic acids 550 to 725 of SEQ ID NO:19, or a complement thereof. The invention features nucleic acid molecules comprising at least 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105, 110, 115, 120, 125, 130, 135, 140, 145, 150, 155, 160, 165, 170, 175, 180, 185, 190, 195, 200, 225, 230, 235, 240, 245, 250, 255, 260, 265, 270, 275, 280, 285, 290, 295, or 300 nucleotides of nucleic acids 1404 to 1713 of SEQ ID NO:19, or a complement thereof.

The invention features nucleic acid molecules comprising at least 420, 425, 450, 475, 500, 525, 550, 600, or 650 nucleotides of the nucleotide sequence of SEQ ID NO:21, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 30, 35, 40, 45, 50, 55 or 60 nucleotides of nucleic acids 1 to 132, or of nucleic acids 549 to 651, of SEQ ID NO:21, or a complement thereof.

The invention features nucleic acid molecules which include a fragment of at least 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1350, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, or 1682 nucleotides of the nucleotide sequence of SEQ ID NO:7, the nucleotide sequence of the human TANGO 262 cDNA clone of ATCC™ Accession No. 207176, or a complement thereof. The invention features nucleic acid molecules comprising at least 25, 50, 100, 150, 200, 250, 300, 350, 400, or 440 nucleotides of nucleic acids 1 to 441 of SEQ ID NO:7, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 525, or 530 nucleotides of nucleic acids 795 to 1329 of SEQ ID NO:7, the nucleotide sequence of the human TANGO 262 cDNA clone of ATCC™ Accession No. 207176, or a complement thereof.

The invention features nucleic acid molecules of at least 355, 340, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, or 677 nucleotides of the nucleotide sequence of SEQ ID NO:9, the nucleotide sequence of a human TANGO 262 cDNA, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 30, 40, 50, 60, 70, 80, 90, 100, 105, 110 or 115 nucleotides of nucleic acids 1 to 120 of SEQ ID NO:9, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 75, 100, 125, 150, 175, or 200 nucleotides of nucleic acids 474 to 678 of SEQ ID NO:9, or a complement thereof.

The invention features nucleic acid molecules of at least 510, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400, or 1425 nucleotides of the nucleotide sequence of SEQ ID NO:22, or a complement thereof..

The invention also features nucleic acid molecules which include a fragment of at least 510, 550, 600, 650, or 677 nucleotides of the nucleotide sequence of SEQ ID NO:24, or a complement thereof.

The invention features nucleic acid molecules of at least 340, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1050, 1100, 1150, 1200, 1250, 1300, 1350, 1400 or 1422 nucleotides of the nucleotide sequence of SEQ ID NO:10, the nucleotide sequence of the human TANGO 266 cDNA clone of ATCC™ Accession No.

207176, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, or 510 nucleotides of nucleic acids 1 to 520 of SEQ ID NO:10, or a complement thereof.

The invention features nucleic acid molecules which include a fragment of at least 25, 50, 75, 100, 125, 150, 175, 200, 225, 250, 275, 300, or 314 nucleotides of the nucleotide sequence of SEQ ID NO:12, or a complement thereof.

The invention features nucleic acid molecules of at least 590, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2250, 2300, 2350, 2400, 2450, 2500, 2550, 2600, 2650, 2700, 2750, 2800, 2850, 2900, or 2925 nucleotides of the nucleotide sequence of SEQ ID NO:13, the nucleotide sequence of the human TANGO 266 cDNA clone of ATCC™ Accession No. 207176, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, 1775, 1800, 1825, 1850, 1875, 1900, or 1925 nucleotides of nucleic acids 1 to 1940 of SEQ ID NO:13, or a complement thereof.

The invention features nucleic acid molecules which include a fragment of at least 590, 600, 650, 700, 750, 800, 850, 900, 950, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, 2100, 2200, 2250, 2300, or 2333 nucleotides of the nucleotide sequence of SEQ ID NO:15, or a complement thereof. The invention also features nucleic acid molecules comprising at least 25, 50, 100, 125, 150, 175, 200, 225, 250, 275, 300, 325, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 725, 750, 775, 800, 825, 850, 875, 900, 925, 950, 975, 1000, 1025, 1050, 1075, 1100, 1125, 1150, 1175, 1200, 1225, 1250, 1275, 1300, 1325, 1375, 1400, 1425, 1450, 1475, 1500, 1525, 1550, 1575, 1600, 1625, 1650, 1675, 1700, 1725, 1750, or 1775 nucleotides of nucleic acids 1 to 1780 of SEQ ID NO:15, or a complement thereof.

The invention features nucleic acid molecules which include a fragment of at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1200) nucleotides of the nucleotide sequence of any of SEQ ID Nos: NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98,100, or 102, or the nucleotide sequence of the cDNA of a clone deposited with ATCC™ as Accession Number 207176, or a complement thereof, wherein

such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention.

The invention also features nucleic acid molecules which include a nucleotide sequence encoding a protein having an amino acid sequence that is at least 45% (or 55%, 65%, 75%, 85%, 95%, or 98%) identical to the amino acid sequence of any of SEQ ID NOs:2, 5, 20, 14, 17, or the amino acid sequence encoded by the cDNA of a clone deposited as ATCC™ as Accession Number 207176, or a complement thereof.

The invention also features nucleic acid molecules which include a nucleotide sequence encoding a protein having an amino acid sequence that is at least 82% (or 83%, 85%, 90%, 92%, 94%, 96%, or 98%) identical to the amino acid sequence of any of SEQ ID NOs:11, or the amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof.

The invention also features nucleic acid molecules which include a nucleotide sequence encoding a protein having an amino acid sequence that is at least 45% (or 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 98%) identical to the amino acid sequence of any of SEQ ID NOs:2, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof, wherein the protein encoded by the nucleotide sequence also exhibits at least one structural and/or functional feature of a polypeptide of the invention.

The invention also features nucleic acid molecules which include a nucleotide sequence encoding a protein having an amino acid sequence that is at least 55% (or 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, or 98%) identical to the amino acid sequence of any of SEQ ID NOs:5, 8, 20, 23, or the amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof, wherein the protein encoded by the nucleotide sequence also exhibits at least one structural and/or functional feature of a polypeptide of the invention.

In preferred embodiments, the nucleic acid molecules have the nucleotide sequence of any of SEQ ID NOs:1, 3, 16, 18 or the nucleotide sequence of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof. Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of any of SEQ ID NOs:2, 17, the fragment including at least 15 (20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, 400, 410, 420, 440, 460, or 475) contiguous amino

acids of any of SEQ ID NOs:2, 17, or the polypeptide encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176.

In preferred embodiments, the nucleic acid molecules have the nucleotide sequence of any of SEQ ID NOs:4, 6, or the nucleotide sequence of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof. Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of any of SEQ ID NOs:5, the fragment including at least 110 (115, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, or 250) contiguous amino acids of any of SEQ ID NOs:5, or the polypeptide encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176.

In preferred embodiments, the nucleic acid molecules have the nucleotide sequence of any of SEQ ID NOs:19, 21, or the nucleotide sequence of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof. Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of any of SEQ ID NOs:20, the fragment including at least 90 (95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, or 215) contiguous amino acids of any of SEQ ID NOs:20, or the polypeptide encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176.

In preferred embodiments, the nucleic acid molecules have the nucleotide sequence of any of SEQ ID NOs:10, 12, or the nucleotide sequence of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof. Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of any of SEQ ID NOs:11, the fragment including at least 85 (86, 88, 90, 92, 94, 96, 98, 100, 102, or 104) contiguous amino acids of any of SEQ ID NOs:11, or the polypeptide encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176.

Also within the invention are nucleic acid molecules which encode a fragment of a polypeptide having the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, the fragment including at least 15 (20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 210, 220, 230, 240, 250, 260, 270, 280, 290, 300, 310, 320, 330, 340, 350, 360, 370, 380, 390, or 400) contiguous amino acids of any of SEQ ID Nos:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or the polypeptide encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, wherein the

fragment exhibits at least one structural and/or functional feature of a polypeptide of the invention.

The invention includes nucleic acid molecules which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof, wherein the nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule having a nucleic acid sequence encoding any of SEQ ID Nos: or a complement thereof.

The invention includes nucleic acid molecules which encode a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof, wherein the nucleic acid molecule hybridizes under stringent conditions to a nucleic acid molecule having a nucleic acid sequence encoding any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or a complement thereof, wherein such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention.

Also within the invention are isolated polypeptides or proteins having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of any of SEQ ID NOs:2, 5, 14, 17, 20.

Also within the invention are isolated polypeptides or proteins having an amino acid sequence that is at least about 82%, preferably 84%, 86%, 88%, 90%, 92%, 94%, 96%, or 98% identical to the amino acid sequence of any of SEQ ID NOs:11.

Also within the invention are isolated polypeptides or proteins having an amino acid sequence that is at least about 65%, preferably 75%, 85%, 95%, or 98% identical to the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the polypeptides or proteins also exhibit at least one structural and/or functional feature of a polypeptide of the invention.

Also within the invention are isolated polypeptides or proteins which preferably are encoded by a nucleic acid molecule having a nucleotide sequence that is at least about 65%, 75%, 85%. or 95% identical the nucleic acid sequence encoding any of SEQ ID NOs:2, 5, 8,

11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the polypeptides or proteins preferably also exhibit at least one structural and/or functional feature of a polypeptide of the invention, and isolated polypeptides or proteins which are encoded by a nucleic acid molecule having a nucleotide sequence which hybridizes under stringent hybridization conditions to a nucleic acid molecule having the sequence of any of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or a complement thereof, or the non-coding strand of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176.

Also within the invention are polypeptides which are naturally occurring allelic variants of a polypeptide that includes the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes under stringent conditions to a nucleic acid molecule having the sequence of any of SEQ ID Nos:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or a complement thereof.

Also within the invention are polypeptides which are naturally occurring allelic variants of a polypeptide that includes the amino acid sequence of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes under stringent conditions to a nucleic acid molecule having the sequence of any of SEQ ID Nos:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or a complement thereof, wherein such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention.

The invention also features nucleic acid molecules that hybridize under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NOs:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof,

wherein preferably such nucleic acid molecules encode polypeptides or proteins that exhibit at least one structural and/or functional feature of a polypeptide of the invention. In other embodiments, the nucleic acid molecules are at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides in length and hybridize under stringent conditions to a nucleic acid molecule comprising the nucleotide sequence of any of SEQ ID NOs:1, 3, 4, 6, 8, 10, 11, 13, 14, 16, 17, 19, 20, 22, 38, 39, 53, 55, 56, 58, 59, 61, 62, 64, 65, 67, 68, 70, 71, 73, of the cDNA of a clone deposited with the ATCC™ as Accession Number 207176, or a complement thereof.

In preferred embodiments, the isolated nucleic acid molecules encode a cytoplasmic, transmembrane, or extracellular domain of a polypeptide of the invention.

In another embodiment, the invention provides an isolated nucleic acid molecule which is antisense to the coding strand of a nucleic acid of the invention.

Another aspect of the invention provides vectors, e.g., recombinant expression vectors, comprising a nucleic acid molecule of the invention. In another embodiment the invention provides host cells containing such a vector. The invention also provides methods for producing a polypeptide of the invention by culturing, in a suitable medium, a host cell of the invention containing a recombinant expression vector encoding a polypeptide of the invention such that the polypeptide of the invention is produced.

Another aspect of this invention features isolated or recombinant proteins and polypeptides of the invention. Preferred proteins and polypeptides possess at least one biological activity possessed by the corresponding naturally-occurring human polypeptide. An activity, a biological activity, and a functional activity of a polypeptide of the invention refers to an activity exerted by a protein or polypeptide of the invention on a responsive cell as determined *in vivo*, or *in vitro*, according to standard techniques. Such activities can be a direct activity, such as an association with or an enzymatic activity on a second protein or an indirect activity, such as a cellular signaling activity mediated by interaction of the protein with a second protein. Thus, such activities include, e.g., (1) the ability to form protein-protein interactions with proteins in the signaling pathway of the naturally-occurring polypeptide; (2) the ability to bind a ligand of the naturally-occurring polypeptide; (3) the ability to bind to an intracellular target of the naturally-occurring polypeptide. Other activities include: (1) the ability to modulate cellular proliferation; (2) the ability to modulate cellular differentiation; (3) the ability to modulate chemotaxis and/or migration; and/or (4) the ability to modulate cell death.

In one embodiment, a polypeptide of the invention has an amino acid sequence sufficiently identical to an identified domain of a polypeptide of the invention. As used

herein, the term "sufficiently identical" refers to a first amino acid or nucleotide sequence which contains a sufficient or minimum number of identical or equivalent (e.g., with a similar side chain) amino acid residues or nucleotides to a second amino acid or nucleotide sequence such that the first and second amino acid or nucleotide sequences have a common structural domain and/or common functional activity. For example, amino acid or nucleotide sequences which contain a common structural domain having about 65% identity, preferably 75% identity, more preferably 85%, 95%, or 98% identity are defined herein as sufficiently identical.

In one embodiment, the isolated polypeptide of the invention lacks both a transmembrane and a cytoplasmic domain. In another embodiment, the polypeptide lacks both a transmembrane domain and a cytoplasmic domain and is soluble under physiological conditions.

The polypeptides of the present invention, or biologically active portions thereof, can be operably linked to a heterologous amino acid sequence to form fusion proteins. The invention further features antibodies that specifically bind a polypeptide of the invention such as monoclonal or polyclonal antibodies. In addition, the polypeptides of the invention or biologically active portions thereof can be incorporated into pharmaceutical compositions, which optionally include pharmaceutically acceptable carriers.

In another aspect, the present invention provides methods for detecting the presence of the activity or expression of a polypeptide of the invention in a biological sample by contacting the biological sample with an agent capable of detecting an indicator of activity such that the presence of activity is detected in the biological sample.

In another aspect, the invention provides methods for modulating activity of a polypeptide of the invention comprising contacting a cell with an agent that modulates (inhibits or stimulates) the activity or expression of a polypeptide of the invention such that activity or expression in the cell is modulated. In one embodiment, the agent is an antibody that specifically binds to a polypeptide of the invention.

In another embodiment, the agent modulates expression of a polypeptide of the invention by modulating transcription, splicing, or translation of an mRNA encoding a polypeptide of the invention. In yet another embodiment, the agent is a nucleic acid molecule having a nucleotide sequence that is antisense to the coding strand of an mRNA encoding a polypeptide of the invention.

The present invention also provides methods to treat a subject having a disorder characterized by aberrant activity of a polypeptide of the invention or aberrant expression of a nucleic acid of the invention by administering an agent which is a modulator of the

activity of a polypeptide of the invention or a modulator of the expression of a nucleic acid of the invention to the subject. In one embodiment, the modulator is a protein of the invention. In another embodiment, the modulator is a nucleic acid of the invention. In other embodiments, the modulator is a peptide, peptidomimetic, or other small molecule.

The present invention also provides diagnostic assays for identifying the presence or absence of a genetic lesion or mutation characterized by at least one of: (i) aberrant modification or mutation of a gene encoding a polypeptide of the invention, (ii) mis-regulation of a gene encoding a polypeptide of the invention, and (iii) aberrant post-translational modification of a polypeptide of the invention wherein a wild-type form of the gene encodes a polypeptide having the activity of the polypeptide of the invention.

In another aspect, the invention provides a method for identifying a compound that binds to or modulates the activity of a polypeptide of the invention. In general, such methods entail measuring a biological activity of the polypeptide in the presence and absence of a test compound and identifying those compounds which alter the activity of the polypeptide.

The invention also features methods for identifying a compound which modulates the expression of a polypeptide or nucleic acid of the invention by measuring the expression of the polypeptide or nucleic acid in the presence and absence of the compound.

In yet a further aspect, the invention provides substantially purified antibodies or fragments thereof, including non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of SEQ ID NO: 2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103 or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of SEQ ID NO: 2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103; an amino acid sequence which is at least 95% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, under conditions of

hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. In various embodiments, the substantially purified antibodies of the invention, or fragments thereof, can be human, non-human, chimeric and/or humanized antibodies.

In another aspect, the invention provides non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103; an amino acid sequence which is at least 95% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. Such non-human antibodies can be goat, mouse, sheep, horse, chicken, rabbit, or rat antibodies. Alternatively, the non-human antibodies of the invention can be chimeric and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies.

In still a further aspect, the invention provides monoclonal antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103; an amino acid sequence which is at least 95% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is

determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

In a particularly preferred embodiment, the substantially purified antibodies or fragments thereof, the non-human antibodies or fragments thereof, and/or the monoclonal antibodies or fragments thereof, of the invention specifically bind to an extracellular domain of the amino acid sequence of SEQ ID NO:27, 39, 41, 46, 47, 51, 122, or 123. Preferably, the extracellular domain to which the antibody, or fragment thereof, binds comprises amino acid residues 34-317 of SEQ ID NO:27, amino acid residues 342-488 of SEQ ID NO:39, amino acid residues 98-317 of SEQ ID NO:41, amino acid residues 34-79 of SEQ ID NO:46, amino acid residues 342-487 of SEQ ID NO:47, amino acid residues 98-317 of SEQ ID NO:51, amino acid residues 1-558 of SEQ ID NO:122, or amino acid residues 773-778 of SEQ ID NO:123.

Any of the antibodies of the invention can be conjugated to a therapeutic moiety or to a detectable substance. Non-limiting examples of detectable substances that can be conjugated to the antibodies of the invention are an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, and a radioactive material.

The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In preferred embodiments, the pharmaceutical composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

Other features and advantages of the invention will be apparent from the following detailed description and claims.

Brief Description of the Drawings

Figure 1A-1C depicts the cDNA sequence of human TANGO 216 (SEQ ID NO:1) and predicted amino acid sequence of human TANGO 216 (SEQ ID NO:2). The open

reading frame of SEQ ID NO:1 extends from nucleotide 307 to 1770 of SEQ ID NO:1 (SEQ ID NO:3).

Figure 2A-2B depicts the cDNA sequence of murine TANGO 216 (SEQ ID NO:16) and predicted amino acid sequence of murine TANGO 216 (SEQ ID NO:17). The open reading frame of SEQ ID NO:16 extends from nucleotide 149 to 1609 of SEQ ID NO:16 (SEQ ID NO:18).

Figure 3 depicts a hydropathy plot of human TANGO 216. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 4 depicts the alignment of the amino acid sequence of human TANGO 216 (SEQ ID NO:2) and murine TANGO 216 (SEQ ID NO:17). In this alignment, a (|) between the two sequences indicates an exact match.

Figure 5 depicts the cDNA sequence of human TANGO 261 (SEQ ID NO:4) and predicted amino acid sequence of human TANGO 261 (SEQ ID NO:5). The open reading frame of SEQ ID NO:4 extends from nucleotide 6 to 761 of SEQ ID NO:4 (SEQ ID NO:6).

Figure 6A-6B depicts the cDNA sequence of a murine TANGO 261 clone (SEQ ID NO:19) and predicted amino acid sequence of murine TANGO 261 (SEQ ID NO:20). The open reading frame of SEQ ID NO:19 extends from nucleotide 2 to 652 of SEQ ID NO:19 (SEQ ID NO:21).

Figure 7 depicts a hydropathy plot of human TANGO 261. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Figure 8 depicts the alignment of the amino acid sequence of human TANGO 261 (SEQ ID NO:5) and a portion of murine TANGO 261 (SEQ ID NO:20). In this alignment, a (|) between the two sequences indicates an exact match.

Figure 9A-9B depicts the cDNA sequence of human TANGO 262 (SEQ ID NO:7) and predicted amino acid sequence of human TANGO 262 (SEQ ID NO:8). The open reading frame of SEQ ID NO:7 extends from nucleotide 322 to 999 of SEQ ID NO:7 (SEQ ID NO:9).

Figure 10 depicts the cDNA sequence of murine TANGO 262 (SEQ ID NO:22) and predicted amino acid sequence of murine TANGO 262 (SEQ ID NO:23). The open reading frame of SEQ ID NO:22 extends from nucleotide 89 to 766 of SEQ ID NO:22 (SEQ ID NO:24).

Figure 11 depicts a hydropathy plot of human TANGO 262. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) are indicated by short vertical lines just below the hydropathy trace.

Figure 12 depicts the alignment of the amino acid sequence of human TANGO 262 (SEQ ID NO:8) and murine TANGO 262 (SEQ ID NO:23). In this alignment, a (|) between the two sequences indicates an exact match.

Figure 13 depicts the alignment of the amino acid sequence of human TANGO 262 (SEQ ID NO:8) and K10C3.4 (SEQ ID NO:38). In this alignment, a (•) between the two sequences indicates an exact match.

Figure 14 depicts the cDNA sequence of human TANGO 266 (SEQ ID NO:10) and predicted amino acid sequence of human TANGO 266 (SEQ ID NO:11). The open reading frame of SEQ ID NO:10 extends from nucleotide 49 to 363 of SEQ ID NO:10 (SEQ ID NO:12).

Figure 15 depicts a hydropathy plot of a human TANGO 266. Relatively hydrophobic residues are above the dashed horizontal line, and relatively hydrophilic residues are below the dashed horizontal line. The cysteine residues (cys) are indicated by short vertical lines just below the hydropathy trace.

Figure 16 depicts the alignment of the amino acid sequence of human TANGO 266 (SEQ ID NO:11) and *Dendroaspis polypepis polypepis* venom protein A (SwissProt Accession Number P25687; SEQ ID NO:36). In this alignment, a (•) between the two sequences indicates an exact match.

Figure 17A-17C depicts the cDNA sequence of human TANGO 267 (SEQ ID NO:13) and predicted amino acid sequence of human TANGO 267 (SEQ ID NO:14). The open reading frame of SEQ ID NO:13 extends from nucleotide 161 to 2494 of SEQ ID NO:13 (SEQ ID NO:15).

Figure 18A-18B depicts the alignment of the amino acid sequence of human TANGO 267 (SEQ ID NO:14) and hepatocellular carcinoma associated gene JCL-1 (GenBank Accession Number U92544; SEQ ID NO:37). In this alignment, a (•) between the two sequences indicates an exact match.

Figure 19A-19D Alignment of Mbkn with snake venom protein VPRA and frog skin secreted protein Bv8, and Western blot showing Mbkn is a secreted protein. A: Amino acid sequence alignment of Mbkn with Bv8 and VPRA. Regions with significant identity are boxed. Numbers correspond to the sequence of the adjacent protein. mBv8-3 is mouse splice variant 3 of Bv8, fBv8 is frog Bv8. B: Schematic diagram with relative phylogenetic

relationship between Mbkn, Bv8, and VPRA. C: Hydrophobicity profile and location of cycteines (cys) of Mbkn. The vertical line represents signal peptide cleaveage site. D: Western Blot analysis of recombinant MbknFc and MbknAP fusion proteins as well as supernatants from 293 cells and 3T3 cell supernatant using affinity purified rabbit anti-Mbkn polyclonal antibodies.

Figure 20A-20B Distribution of Mbkn expression in endocrine organs. A: Northern blot analysis of multiple human tissue RNAs hybridized with a Mbkn probe. B: Relative expression of Mbkn in multiple human tissues by quantitative PCR of cDNA
C: In situ: High expression detected in the ovarian stroma, but no expression was detected in the ovarian endothelium. Moderate expression detected in the placenta

Figure 21 Mbkn binds to mouse macrophages. Mbkn-AP specifically binds to cultured mouse macrophages using Mbkn-AP fusion protein, and is inhibited from binding by Mbkn-Fc fusion protein.

Description of the Preferred Embodiments

The present invention is based on the discovery of cDNA molecules encoding TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267, all of which are either wholly secreted or transmembrane proteins.

TANGO 216

In one aspect, the present invention is based on the discovery of cDNA molecules which encode a novel family of proteins having a von Willebrand factor (vWF) A domain, referred to herein as TANGO 216 proteins. Described herein are human TANGO 216 (SEQ ID NO:1), and murine TANGO 216 (SEQ ID NO:16) nucleic acid molecules and the corresponding polypeptides which the nucleic acid molecules encode (SEQ ID NO:2 and SEQ ID NO:17, respectively).

The TANGO 216 proteins and nucleic acid molecules comprise a family of molecules having certain conserved structural and functional features. As used herein, the term "family" is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Family members can be from either the same or different species. For example, a family can comprises two or more proteins of human origin, or can comprise one or more proteins of human origin and one or more of non-human origin. Members of the same family may also have common structural domains.

For example, the TANGO 216 proteins of the invention include a domain which bears sequence identity to a vWF A domain. Proteins having such a domain are involved in biological processes controlled by specific, often adhesive, molecular interactions. The vWF A domain mediates binding to proteins and sugars. Proteins having vWF A domains may interact through homophilic interactions between vWF A domains. Thus, included within the scope of the invention are TANGO 216 proteins having a vWF A domain. As used herein, a vWF A domain refers to an amino acid sequence of about 150 to 190, preferably about 155 to 185, 160 to 180, and more preferably about 170 amino acids in length. Conserved amino acid motifs, referred to herein as "consensus patterns" or "signature patterns", can be used to identify TANGO 216 family members. For example, the following signature pattern can be used to identify TANGO 216 family members: D - x (2) - F -[ILV] - x - D - x - S - x (2, 3) - [ILV]- x (10, 12) - F. The signature patterns or consensus patterns described herein are described according to the following designation: all amino acids are indicated according to their universal single letter designation; "x" designates any amino acid; x(n) designates "n" number of amino acids, e.g., x (2) designates any two amino acids, e.g., x (2, 3) designates any of two to three amino acids; and, amino acids in brackets indicates any one of the amino acids within the brackets, e.g., [ILV] indicates any of one of either I (isoleucine), L (leucine) or V (valine). TANGO 216 has such a signature pattern at about amino acids 44 to 169 of SEQ ID NO:2 (SEQ ID NO:).

The vWF A domain consensus sequence is also available from the HMMer version 2.0 software as Accession Number PF00092 (SEQ ID NO:). Software for HMM-based profiles is available from http://www.csc.ucsc.edu/research/compbio/sam.html and from http://genome.wustl.edu/eddy/hmmer.html. A vWF A domain of TANGO 216 extends, for example, from about amino acids 44 to 213 of SEQ ID NO:2 (SEQ ID NO:).

Also included within the scope of the present invention are TANGO 216 proteins having a signal sequence. As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 15 to 40 amino acid residues, preferably about 15-30 amino acid residues, and has at least about 60-80%, more preferably 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer.

In certain embodiments, a TANGO 216 family member has the amino acid sequence of SEQ ID NO:2, and the signal sequence is located at amino acids 1 to 31, 1 to 32, 1 to 33, 1 to 34 or 1 to 35. In such embodiments of the invention, the domains and the mature protein resulting from cleavage of such signal peptides are also included herein. For example, the cleavage of a signal sequence consisting of amino acids 1 to 33 of SEQ ID NO:2 (SEQ ID NO:25) results in a mature TANGO 216 protein corresponding to amino acids 34 to 488 of SEQ ID NO:2 (SEQ ID NO:26). The signal sequence is normally cleaved during processing of the mature protein.

The present invention also includes TANGO 216 proteins having a transmembrane domain. As used herein, a transmembrane domain refers to an amino acid sequence having at least about 25 to 40 amino acid residues in length and which contains hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a transmembrane domain contains at least about 25 to 40 amino acid residues, preferably about 25-30 amino acid residues, and has at least about 60-80% hydrophobic residues. An example of a transmembrane domain includes from about amino acids 318 to 345 of SEQ ID NO:2 (SEQ ID NO: 28).

In one embodiment, a TANGO 216 protein of the invention includes a vWF A domain. In another embodiment, a TANGO 216 protein of the invention includes a vWF A domain, and a signal sequence. In another embodiment, a TANGO 216 protein of the invention includes a vWF A domain, a extracellular domain, and a signal sequence. In another embodiment, a TANGO 216 protein of the invention includes a vWF A domain, and an extracellular domain. In another embodiment, a TANGO 216 protein of the invention includes a vWF A domain, an extracellular domain, and a transmembrane domain. In another embodiment, a TANGO 216 protein of the invention includes a vWF A domain, an extracellular domain, a transmembrane domain, and a cytoplasmic domain.

Various features of human and mouse TANGO 216 are summarized below.

Human TANGO 216

The cDNA encoding human TANGO 216 was isolated by screening for cDNAs which encode a potential signal sequence. Briefly, a clone encoding TANGO 216 was isolated through high throughput screening of a prostate stroma cell library. The human TANGO 216 clone includes a 3677 nucleotide cDNA (*Figure 1A-1C*; SEQ ID NO:1). The open reading frame of this cDNA, nucleotides 307 to 1770 (SEQ ID NO:3), encodes a 488 amino acid transmembrane protein depicted in *Figure 1A-1C* (SEQ ID NO:2).

In another embodiment, a human TANGO 216 clone comprises a 4350 nucleotide

cDNA (SEQ ID NO:127). The open reading frame of this cDNA comprises nucleotides 353 to 1819 (SEQ ID NO:128), and encodes a transmembrane protein comprising the 488 amino acid sequence depicted in SEQ ID NO:2.

In one embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 318 is a guanine (G)(SEQ ID NO:1). In this embodiment, the amino acid at position 12 is glutamate (E)(SEQ ID NO:2). In another embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 318 is a cytosine (C)(SEQ ID NO:56). In this embodiment, the amino acid at position 12 is aspartate (D)(SEQ ID NO:57) In another embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 411 is a guanine (G)(SEQ ID NO:1). In this embodiment, the amino acid at position 35 is a glutamate (E)(SEQ ID NO:2). In another embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 411 is a cytosine (C)(SEQ ID NO:58). In this embodiment, the amino acid at position 35 is aspartate (D)(SEQ ID NO:59). In another embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 489 is an adenine (A)(SEQ ID NO:1). In this embodiment, the amino acid at position 61 is a glutamate (E)(SEQ ID NO:2). In another embodiment of a nucleotide sequence of human TANGO 216, the nucleotide at position 489 is a cytosine (C)(SEQ ID NO:60). In this embodiment, the amino acid at position 61 is aspartate (D)(SEQ ID NO:61).

The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that human TANGO 216 includes a 33 amino acid signal peptide (amino acids 1 to about amino acid 33 of SEQ ID NO:2; SEQ ID NO: 25) preceding the mature TANGO 216 protein (corresponding to about amino acid 34 to amino acid 488 of SEQ ID NO:2; SEQ ID NO:26). The presence of a methionine residue at positions 78, 245, 277, 337, 392, and 369 of SEQ ID NO:2 indicate that there can be alternative forms of human TANGO 216 of 411 amino acids of SEQ ID NO:2, 244 amino acids of SEQ ID NO:2, 212 amino acids of SEQ ID NO:2, 152 amino acids of SEQ ID NO:2, 97 amino acids of SEQ ID NO:2, and 120 amino acids of SEQ ID NO:2, respectively.

In one embodiment, human TANGO 216 includes extracellular domains (about amino acids 34 to 79 and 342 to 488 of SEQ ID NO:2)(SEQ ID NO:27 and SEQ ID NO:39, respectively), transmembrane (TM) domains (amino acids 80-97 and 318 to 341 of SEQ ID NO:2)(SEQ ID NO:28, and SEQ ID NO:40, respectively); and a cytoplasmic domain (amino acids 98 to 317 of SEQ ID NO:2)(SEQ ID NO:29). The cytoplasmic domain is very rich in proline and glutamic acid residues. These residues represent 27% of the residues in the cytoplasmic domain of human TANGO 216.

Alternatively, in another embodiment, a human TANGO 216 protein contains an extracellular domain at amino acid residues 98 to 317 of SEQ ID NO:2 (SEQ ID NO:41), transmembrane (TM) domains (amino acids 80-97 and 318 to 341 of SEQ ID NO:2 (SEQ ID NO:42, and SEQ ID NO:43, respectively), and cytoplasmic domains at amino acid residues 1 to 79 and 342-488 of SEQ ID NO:2 (SEQ ID NO:44 and SEQ ID NO:45, respectively).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the human TANGO 216 amino acid sequence in SEQ ID NO:2, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of human TANGO 216, nucleotides 310-1770 (SEQ ID NO:106), encodes the human TANGO 216 amino acid sequence from amino acids 2-488 (SEQ ID NO:107).

Human TANGO 216 includes a vWF A domain from about amino acids 44 to 213 of SEQ ID NO:2 (SEQ ID NO:37).

Human TANGO 216 protein, including the signal sequence, has a molecular weight of 53.6 kDa prior to post-translational modification. Human TANGO 216 protein has a molecular weight of 50.0 kDa after cleavage of the 33 amino acid signal peptide.

A clone, EpT216, which encodes human TANGO 216 was deposited with the American Type Culture Collection (ATCC™, 10801 University Boulevard, Manassas, VA 20110-2209) on March 26, 1999, and was assigned Accession Number 207176. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience to those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Figure 3 depicts a hydropathy plot of human TANGO 216. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. As shown in the hydropathy plot, the hydrophobic region at the beginning of the plot which corresponds to about amino acids 1 to 33 of SEQ ID NO:2 is the signal sequence of TANGO 216 (SEQ ID NO:). The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Northern analysis of human TANGO 216 mRNA expression revealed the presence of an approximately 3.8 kb transcript and an approximately 4.3 kb transcript that are expressed in a range of tissues including lung, liver, skeletal muscle, kidney, and pancreas,

with highest expression in heart and placenta. The two transcripts likely represent alternative poly A site usage.

The human gene for TANGO 216 was mapped on radiation hybrid panels to the long arm of chromosome 4, in the region q11-13. Flanking markers for this region are GCT14E02 and jktbp-rs2. The JPD (periodontitis, juvenile), and DGI1(dentinogenesis imperfecta) loci also map to this region of the human chromosome. The GRO1 (FRO1 oncogene), ALB (albumin), IL8 (interleukin 8), HTN (histatin), and DCK (deoxycytidine kinase) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 5. The rs (recessive spotting) locus also maps to this region of the mouse chromosome. The ste (sulfotransferase), areg (amphiregulin), btc (betacellulin), mc (marcel), alb1 (albumin 1), and afp (alpha fetoprotein) genes also map to this region of the mouse chromosome.

Mouse TANGO 216

A mouse homolog of human TANGO 216 was identified. A cDNA encoding mouse TANGO 216 was identified by analyzing the sequences of clones present in a mouse bone marrow cDNA library. This analysis led to the identification of a clone, jtmMa005g09, encoding mouse TANGO 216. The murine TANGO 216 cDNA of this clone is 3501 nucleotides long (Figure 2A-2B; SEQ ID NO:16). The open reading frame of this cDNA comprises nucleotides 149 to 1609 of SEQ ID NO:16 (SEQ ID NO:18) and encodes the 487 amino acid protein depicted in Figure 2A-2B (SEQ ID NO:17).

In another embodiment, a mouse TANGO 216 clone includes comprises a 3647 nucleotide cDNA (SEQ ID NO:129). The open reading frame of this cDNA comprises nucleotides 32 to 469 (SEQ ID NO:131), and encodes a transmembrane protein comprising the 146 amino acid sequence depicted in SEQ ID NO:130.

In one embodiment, mouse TANGO 216 includes extracellular domains (about amino acids 34 to 79 and 342 to 487 of SEQ ID NO:17 (SEQ ID NO:46 and SEQ ID NO:47, respectively), transmembrane (TM) domains (amino acids 80-97 and 318 to 341 of SEQ ID NO:17)(SEQ ID NO:48 and SEQ ID NO:49, respectively); and a cytoplasmic domain (amino acids 98 to 317 of SEQ ID NO:17)(SEQ ID NO:50). The cytoplasmic domain is very rich in proline and glutamic acid residues. These residues represent 27% of the residues in the cytoplasmic domain of human TANGO 216. Alternatively, in another embodiment, a mouse TANGO 216 protein contains an extracellular domain at amino acid residues 98 to 317 of SEQ ID NO:17 (SEQ ID NO:51), transmembrane (TM) domains (amino acids 80-97 and 318 to 341 of SEQ ID NO:17 (SEQ ID NO:52 and SEQ ID NO:53,

respectively), and cytoplasmic domains at amino acid residues 1 to 79 and 342-487 of SEQ ID NO:17 (SEQ ID NO:54 and SEQ ID NO:55, respectively).

The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that mouse TANGO 216 includes a 33 amino acid signal peptide (amino acids 1 to about amino acid 33 of SEQ ID NO:17; SEQ ID NO:146) preceding the mature TANGO 216 protein (corresponding to about amino acid 34 to amino acid 48 of SEQ ID NO:17; SEQ ID NO:147). The presence of a methionine residue at positions 78, 337, 360, 392, 417, 459, and 468 of SEQ ID NO:17 indicate that there can be alternative forms of mouse TANGO 216 of 410 amino acids of SEQ ID NO:17, 151 amino acids of SEQ ID NO:17, 128 amino acids of SEQ ID NO:17, 96 amino acids of SEQ ID NO:17, 71 amino acids of SEQ ID NO:17, 29 amino acids of SEQ ID NO:17, and 20 amino acids of SEQ ID NO:17, respectively.

In one embodiment of a nucleotide sequence of mouse TANGO 216 the nucleotide at position 253 is a guanine (G)(SEQ ID NO:16). In this embodiment, the amino acid at position 35 is glutamate (E)(SEQ ID NO:17). In another embodiment of a nucleotide sequence of mouse TANGO 216, the nucleotide at position 253 is a cytosine (C)(SEQ ID NO:62). In this embodiment, the amino acid at position 35 is aspartate (D)(SEQ ID NO:63). In another embodiment of a nucleotide sequence of mouse TANGO 216, the nucleotide at position 331 is an adenine (A)(SEQ ID NO:16). In this embodiment, the amino acid at position 61 is a glutamate (E)(SEQ ID NO:17). In another embodiment of a nucleotide sequence of mouse TANGO 216, the nucleotide at position 331 is a cytosine (C)(SEQ ID NO:64). In this embodiment, the amino acid at position 61 is aspartate (D)(SEQ ID NO:65). In another embodiment of a nucleotide sequence of mouse TANGO 216, the nucleotide at position 371 is a guanine (G)(SEQ ID NO:16). In this embodiment, the amino acid at position 71 is a glutamate (E)(SEQ ID NO:17). In another embodiment of a nucleotide sequence of mouse TANGO 216, the nucleotide at position 371 is a cytosine (C)(SEQ ID NO:66). In this embodiment, the amino acid at position 71 is aspartate (D)(SEQ ID NO:67).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the mouse TANGO 216 amino acid sequence in SEQ ID NO:17, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of mouse TANGO 216, nucleotides 152-1609 (SEQ ID NO:108), encodes the mouse TANGO 216 amino acid sequence comprising amino acids 2-487 (SEQ ID NO:109).

Mouse TANGO 216 includes a vWF A domain from about amino acids 44 to 213 of SEQ ID NO:17 (SEQ ID NO:148).

Mouse TANGO 216 protein, including the signal sequence, has a molecular weight of 53.2 kDa prior to post-translational modification. Mouse TANGO 216 protein has a molecular weight of 49.8 kDa after cleavage of the 33 amino acid signal peptide.

In situ tissue screening was performed on mouse adult and embryonic tissue to analyze the expression of mouse TANGO 216 mRNA. In the case of adult expression, a low level ubiquitous signal was detected in the spleen and stomach. A weak, ubiquitous signal was detected in the thymus. A ubiquitous signal was detected in the liver, submandibular salivary gland, heart, colon, and in the cortical region of the adrenal gland. A multifocal pattern was detected in the lung and in the decidua of the placenta. A signal was apparent in the villi of the small intestine. No signal was detected in the following tissues: brain, spinal cord, eye, brown fat, white fat, pancreas, skeletal muscle, bladder, kidney, and lung.

In the case of embryonic expression, expression was seen in a number of tissues. At E13.5, strong signals were detected in the developing spinal column, heart, and tongue. Meckelís cartilage was also apparent. Limb expression is not readily apparent. Low level signal was also seen throughout the gut region including but not restricted to lung, liver, and intestines. Signal is noticeably absent from the developing CNS except for the areas of the brain surrounding the lateral ventricals and mesencephalic vesicle. At E14.5, developing spinal column and sternum, heart, tongue, and Meckelís cartilage continued to have strong signal. Signal from the heart and tongue was ubiqutious. In the brain, the diencephalon had the strongest signal with the areas surrounding the ventricles still being positive. At E15.5, signal was seen in the previously stated regions and was readily seen in the primordium of the basisphenoid bone and primordium of the nasal bone. At E16.5, signal was seen in the previously stated regions, primordium of the basisphenoid bone. At E18.5, the strongest signal was obtained in the developing bone and cartilage areas. Signal from the heart was diminished in strength and now equal to that seen in the rest of the gut region. At P1.5, signal was still strong in the spinal column and nasal septum. Signal was absent from the CNS except for faint signal in the region of the developing cerebellum. Signal is otherwise low and ubiquitous except for heart, small intestine, and stomach which have a slightly higher signal. The highest expressing tissue was the capsule of the kidney which was seen at E14.5 and continues to P1.5.

Human and murine TANGO 216 sequences exhibit considerable similarity at the protein, nucleic acid, and open reading frame levels. An alignment (made using the ALIGN software (Myers and Miller (1989) CABIOS, ver. 2.0); BLOSUM 62 scoring matrix; gap

penalties -12/-4), reveals a protein identity of 84.8%. The human and murine TANGO 216 full length cDNAs are 84.4% identical, as assessed using the same software and parameters as indicated (without the BLOSUM 62 scoring matrix). In the respective ORFs, calculated in the same fashion as the full length cDNAs, human and murine TANGO 216 are 84% identical.

Figure 4 depicts the alignment of the amino acid sequence of human TANGO 216 (SEQ ID NO:2) and murine TANGO 216 (SEQ ID NO:17). In this alignment, a (|) between the two sequences indicates an exact match. The alignment of the amino acid sequence of human TANGO 216 (SEQ ID NO:2) and murine TANGO 216 (SEQ ID NO:130) over the 146 amino acid sequence of murine TANGO 216 (SEQ ID NO:130) indicated a percent identity of approximately 65-68%.

Uses of TANGO 216 Nucleic Acids, Polypeptides, and Modulators Thereof

The TANGO 216 proteins of the invention include a vWF A domain. Accordingly, TANGO 216 proteins likely function in a similar manner as other proteins which include a vWF A domain, including von Willebrand factor, a large multimeric protein found in platelets, endothelial cells, and plasma. Thus, TANGO 216 modulators can be used to treat any von Willebrand factor-associated disorders and modulate normal von Willebrand factor functions.

As discussed above, the vWF domain of TANGO 216 is involved in cellular adhesion and interaction with extracellular matrix (ECM) components. Proteins of the type A module superfamily which incorporate a vWF domain participate in multiple ECM and cell/ECM interactions. For example, proteins having a vWF domain have been found to play a role in cellular adhesion, migration, homing, pattern formation and/or signal transduction after interaction with several different ligands (Colombatti et al. (1993) *Matrix* 13:297-306).

Similarly, the TANGO 216 proteins of the invention likely play a role in various extracellular matrix interactions, e.g., matrix binding, and/or cellular adhesion. Thus, a TANGO 216 activity is at least one or more of the following activities: 1) regulation of extracellular matrix structuring; 2) modulation of cellular adhesion, either *in vitro* or *in vivo*; 3) regulation of cell trafficking and/or migration. Accordingly, the TANGO 216 proteins, nucleic acid molecules and/or modulators can be used to modulate cellular interactions such as cell-cell and/or cell-matrix interactions and thus, to treat disorders associated with abnormal cellular interactions.

TANGO 216 polypeptides, nucleic acids and/or modulators thereof can also be used to modulate cell adhesion in proliferative disorders, such as cancer. Examples of types of cancers include benign tumors, neoplasms or tumors (such as carcinomas, sarcomas, adenomas or myeloid lymphoma tumors, *e.g.*, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leimyosarcoma, rhabdotheliosarcoma, colon sarcoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadenocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hematoma, bile duct carcinoma, melanoma, choriocarcinoma, semicoma, embryonal carcinoma, Wilms' tumor, cervical cancer, testicular tumor, lung carcinoma, small cell carcinoma, bladdcr carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependynoma, pinealoma, hemangioblastoma, retinoblastoma), leukemias, (*e.g.* acute lymphocytic leukemia), acute myelocytic leukemia (myelolastic, promyelocytic, myelomonocytic, monocytic and erythroleukemia), chronic leukemias (chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia), or polycythemia vera, or lymphomas (Hodgkin's disease and non-Hodgkin's diseases), multiple myelomas and Waldenström's macroglobulinemia.

As TANGO 216 was originally isolated from a bone marrow library, TANGO 216 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of cells that appear in the bone marrow, e.g., stem cells (e.g., hematopoietic stem cells), and blood cells, e.g., erythrocytes, platelets, and leukocytes. Thus TANGO 269 nucleic acids, proteins, and modulators thereof can be used to treat bone marrow, blood, and hematopoietic associated diseases and disorders, e.g., acute myeloid leukemia, hemophilia, leukemia, anemia (e.g., sickle cell anemia), and thalassemia.

As TANGO 216 exhibits expression in the embryonic lung, TANGO 216 polypeptides, nucleic acids, or modulators thereof, can be used to treat pulmonary (lung) disorders, such as atelectasis, pulmonary congestion or edema, chronic obstructive airway disease (e.g., emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis), diffuse interstitial diseases (e.g., sarcoidosis, pneumoconiosis, hypersensitivity pneumonitis, Goodpasture's syndrome, idiopathic pulmonary hemosiderosis, pulmonary alveolar proteinosis, desquamative interstitial pneumonitis, chronic interstitial pneumonia, fibrosing alveolitis, hamman-rich syndrome, pulmonary eosinophilia, diffuse interstitial fibrosis,

Wegener's granulomatosis, lymphomatoid granulomatosis, and lipid pneumonia), or tumors (e.g., bronchogenic carcinoma, bronchiolovlveolar carcinoma, bronchial carcinoid, hamartoma, and mesenchymal tumors).

As TANGO 216 exhibits expression in the small intestine, TANGO 216 polypeptides, nucleic acids, or modulators thereof, can be used to treat intestinal disorders, such as ischemic bowel disease, infective enterocolitis, Crohn's disease, benign tumors, malignant tumors (e.g., argentaffinomas, lymphomas, adenocarcinomas, and sarcomas), malabsorption syndromes (e.g., celiac disease, tropical sprue, Whipple's disease, and abetalipoproteinemia), obstructive lesions, hernias, intestinal adhesions, intussusception, or volvulus.

As TANGO 216 exhibits expression in the spleen, TANGO 216 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, diffirentiation, and/or function of cells that form the spleen, e.g., cells of the splenic connective tissue, e.g., splenic smooth muscle cells and/or endothelial cells of the splenic blood vessels. TANGO 216 nucleic acids, proteins, and modulators thereof can also be used to modulate the proliferation, differentiation, and/or function of cells that are processed, *e.g.*, regenerated or phagocytized within the spleen, e.g., erythrocytes and/or B and T lymphocytes and macrophages. Thus, TANGO 216 nucleic acids, proteins, and modulators thereof can be used to treat spleen, *e.g.*, the fetal spleen, associated diseases and disorders. Examples of splenic diseases and disorders include *e.g.*, splenic lymphoma and/or splenomegaly, and/or phagocytotic disorders, *e.g.*, those inhibiting macrophage engulfment of bacteria and viruses in the bloodstream.

As TANGO 216 is expressed in the kidney, the TANGO 216 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues in which it is expressed. Such molecules can also be used to treat disorders associated with abnormal or aberrant metabolism or function of cells in the tissues in which it is expressed. Such can be used to treat or modulate renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal disease, medullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial

nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atheroembolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

As TANGO 216 exhibits expression in the heart, TANGO 216 polypeptides, nucleic acids, or modulators thereof, can be used to treat cardiovascular disorders, such as ischemic heart disease (*e.g.*, angina pectoris, myocardial infarction, and chronic ischemic heart disease), hypertensive heart disease, pulmonary heart disease, valvular heart disease (*e.g.*, rheumatic fever and rheumatic heart disease, endocarditis, mitral valve prolapse, and aortic valve stenosis), congenital heart disease (*e.g.*, valvular and vascular obstructive lesions, atrial or ventricular septal defect, and patent ductus arteriosus), or myocardial disease (*e.g.*, myocarditis, congestive cardiomyopathy, and hypertrophic cardiomyopathy).

As TANGO 216 exhibits expression in bone structures, TANGO 216 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of bone and cartilage cells, *e.g.*, chondrocytes and osteoblasts, and to treat bone and/or cartilage associated diseases or disorders. Examples of bone and/or cartilage diseases and disorders include bone and/or cartilage injury due to for example, trauma (*e.g.*, bone breakage, cartilage tearing), degeneration (*e.g.*, osteoporosis), degeneration of joints, *e.g.*, arthritis, *e.g.*, osteoarthritis, and bone wearing.

The extracellular region of TANGO 216 has significant similarity to TANGO 197, a secreted protein. TANGO 197 has a vWF A domain and may interact with TANGO 216.

TANGO 216 likely plays a role in the regulation of binding of cells in circulation to the endothelial substrate. Thus, TANGO 216 may regulate proper flow of cells in the heart, vasculature, and placenta. Accordingly, the TANGO 216 proteins, nucleic acids and/or modulators of the invention are useful modulators of interactions between cells in circulation and endothelial substrate which can be used to treat disorders of such interactions.

<u>TANGO 261</u>

In another aspect, the present invention is based on the discovery of nucleic acid sequences which encode a novel family of proteins referred to herein as TANGO 261 proteins. Described herein are human TANGO 261 (SEQ ID NO:4), and murine TANGO 261 (SEQ ID NO:19) nucleic acid molecules and the corresponding polypeptides which the nucleic acid molecules encode (SEQ ID NO:5 and SEQ ID NO:20, respectively).

The TANGO 261 proteins and nucleic acid molecules comprise a family of molecules having certain conserved structural and functional features. As used herein, the term "family" is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Family members can be from either the same or different species. For example, a family can comprises two or more proteins of human origin, or can comprise one or more proteins of human origin and one or more of non-human origin. Members of the same family may also have common structural domains.

Also included within the scope of the present invention are TANGO 261 proteins having a signal sequence. As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 15 to 40 amino acid residues, preferably about 15-30 amino acid residues, and has at least about 60-80%, more preferably 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer.

In certain embodiments, a TANGO 261 family member has the amino acid sequence of SEQ ID NO:5, and the signal sequence is located at amino acids 1 to 26, 1 to 27, 1 to 28, 1 to 29 or 1 to 30. In such embodiments of the invention, the domains and the mature protein resulting from cleavage of such signal peptides are also included herein. For example, the cleavage of a signal sequence consisting of amino acids 1 to 28 of SEQ ID NO:5 (SEQ ID NO:30) results in a mature TANGO 261 protein corresponding to amino acids 29 to 252 of SEQ ID NO:5 (SEQ ID NO:31). The signal sequence is normally cleaved during processing of the mature protein. Thus, in one embodiment, a TANGO 261 protein includes a signal sequence and is secreted.

Various features of human and mouse TANGO 261 are summarized below.

<u>Human TANGO 261</u>

A cDNA clone, jthda088f09, encoding full length human TANGO 261 was identified by screening a stimulated human smooth muscle cell library by EST analysis. Another cDNA clone, jthkf124b08, encoding full length human TANGO 261 was identified by screening a stimulated keratinocyte cell library by EST analysis. The 969 nucleotide human TANGO 261 sequence (SEQ ID NO:4) includes a open reading frame which extends

from nucleotide 6 to nucleotide 761 of SEQ ID NO:4 (SEQ ID NO:6) and encodes a 252 amino acid secreted protein depicted in SEQ ID NO:5.

In another embodiment, a human TANGO 261 clone includes comprises a 1942 nucleotide cDNA (SEQ ID NO:132). The open reading frame of this cDNA comprises nucleotides 146 to 904 (SEQ ID NO:133), and encodes a transmembrane protein comprising the 252 amino acid sequence depicted in Figure 5A-5B (SEQ ID NO:5).

In one embodiment of a nucleotide sequence of human TANGO 261 the nucleotide at position 14 is a guanine (G)(SEQ ID NO:4). In this embodiment, the amino acid at position 3 is glutamate (E)(SEQ ID NO:5). In another embodiment of a nucleotide sequence of human TANGO 261, the nucleotide at position 14 is a cytosine (C)(SEQ ID NO:68). In this embodiment, the amino acid at position 3 is aspartate (D)(SEQ ID NO:69) In another embodiment of a nucleotide sequence of human TANGO 261, the nucleotide at position 149 is an adenine (A)(SEQ ID NO:4). In this embodiment, the amino acid at position 48 is a glutamate (E)(SEQ ID NO:5). In another embodiment of a nucleotide sequence of human TANGO 261, the nucleotide at position 149 is a cytosine (C)(SEQ ID NO:70). In this embodiment, the amino acid at position 48 is aspartate (D)(SEQ ID NO:71). In another embodiment of a nucleotide sequence of human TANGO 261, the nucleotide at position 167 is an adenine (A)(SEQ ID NO:4). In this embodiment, the amino acid at position 54 is a glutamate (E)(SEQ ID NO:5). In another embodiment of a nucleotide sequence of human TANGO 261, the nucleotide at position 167 is a cytosine (C)(SEQ ID NO:72). In this embodiment, the amino acid at position 54 is aspartate (D)(SEQ ID NO:73).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the human TANGO 261 amino acid sequence in SEQ ID NO:5, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of human TANGO 261, nucleotides 9-761 (SEQ ID NO:110), encodes the human TANGO 261 amino acid sequence comprising amino acids 2-252 (SEQ ID NO:111).

Human TANGO 261 includes a signal sequence (amino acid 1 to about amino acid 28 of SEQ ID NO:5; SEQ ID NO:30) preceding the mature protein (about amino acid 29 to amino acid 252 of SEQ ID NO:5; SEQ ID NO:31). The presence of a methionine residue at positions 16, 17, 19, 162, and 190 of SEQ ID NO:5 indicate that there can be alternative forms of human TANGO 261 of 237 amino acids of SEQ ID NO:5, 236 amino acids of SEQ ID NO:5, 234 amino acids of SEQ ID NO:5, 91 amino acids of SEQ ID NO:5, and 63 amino acids of SEQ ID NO:5, respectively.

Human TANGO 261 protein, including the signal sequence, has a molecular weight of 27.9 kD prior to post-translational modification. Mature human TANGO 261 protein has a molecular weight of 24.8 kD prior to post-translational modification.

A clone, EpT261, which encodes human TANGO 261 was deposited with the American Type Culture Collection (ATCC™, 10801 University Boulevard, Manassas, VA 20110-2209) on March 26, 1999, and assigned Accession Number 207176. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Figure 7 depicts a hydropathy plot of human TANGO 261. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. As shown in the hydropathy plot, the hydrophobic region of the plot which corresponds to amino acid 1 to about amino acid 28 of SEQ ID NO:5 is the signal sequence of TANGO 261 (SEQ ID NO:30). The cysteine residues (cys) and potential N-glycosylation sites (Ngly) are indicated by short vertical lines just below the hydropathy trace.

Northern analysis of human TANGO 261 mRNA expression revealed the presence of an approximately 2.6 kb transcript and an approximately 6.0 kb transcript that are expressed in a range of tissues including lung, liver, kidney, and placenta, with highest expression in heart and skeletal muscle. No expression was observed in colon, thymus, peripheral blood leukocytes, and spleen. The two transcripts likely represent alternative poly A site usage.

Human TANGO 261 is likely expressed in prostate epithelium, prostate smooth muscle, bone, and brain, based on the origin of ESTs.

The human gene for TANGO 261 was mapped on radiation hybrid panels to the long arm of chromosome 20, in the region q13.2-13.3 Flanking markers for this region are WI-3773 and AFMA202YB9. The EEGV1 (electroencephalographic variant pattern 1) and PHP1B (pseudohypoparathyroidism) loci also map to this region of the human chromosome. The MC3R (melanocortin 3 receptor), EDN3 (endothelin 3), ADA (adenosine deaminase), and OQTL (obesity QTL) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 2. The fc (flecking) and ra (ragged) loci also map to this region of the mouse chromosome. The mc3r (melanocortin 3 receptor), fc (flecking), ra (ragged), and ntsr (neurotensin receptor) genes also map to this region of the mouse chromosome.

The open reading frame of human TANGO 261 bears significant similarity to the open reading frame of human clone 22 mRNA, alternative splice variant beta 2 (GenBank Accession Number AF009427; Sanders et al. (1997) *Am J. Med. Genet.* 74:140-9), a gene which has brain-specific expression, produces an 8 kb mRNA encoding a 230 amino acid protein, and maps near the candidate region for bipolar affective disorder on chromosome 18. Human TANGO 261 protein and the protein encoded by clone 22 mRNA, alternative splice variant beta 2 are approximately 70% identical. However, human TANGO 261 does not appear to be brain specific.

Mouse TANGO 261

A mouse homolog of human TANGO 261 was identified. A cDNA encoding mouse TANGO 261 was identified by analyzing the sequences of clones present in a mouse microglial cell cDNA library. This analysis led to the identification of a clone, jtmxa004g06, encoding mouse TANGO 261. The murine TANGO 261 cDNA of this clone is 1713 nucleotides long (SEQ ID NO:19). The open reading frame of this cDNA comprises nucleotides 2 to 652 of SEQ ID NO:19 (SEQ ID NO:21) and encodes a protein comprising the 217 amino acid sequence protein depicted in SEQ ID NO:20.

In another embodiment, a mouse TANGO 261 clone includes comprises a 484 nucleotide cDNA (SEQ ID NO:134). The open reading frame of this cDNA comprises nucleotides 3 to 413 (SEQ ID NO:136), and encodes a transmembrane protein comprising the 137 amino acid sequence depicted in Figure 6 (SEQ ID NO:135).

The predicted molecular weight of a mouse TANGO 261 protein without post-translational modifications is 23.9 kDa. The presence of a methionine residue at positions 42, 136, and 160 of SEQ ID NO:20 indicate that there can be alternative forms of mouse TANGO 261 comprising 176 amino acids of SEQ ID NO:20, 82 amino acids of SEQ ID NO:20, and 58 amino acids of SEQ ID NO:20, respectively.

In one embodiment of a nucleotide sequence of mouse TANGO 261 the nucleotide at position 85 is an adenine (A)(SEQ ID NO:19). In this embodiment, the amino acid at position 28 is glutamate (E)(SEQ ID NO:20). In another embodiment of a nucleotide sequence of mouse TANGO 261, the nucleotide at position 85 is a cytosine (C)(SEQ ID NO:74). In this embodiment, the amino acid at position 28 is aspartate (D)(SEQ ID NO:75). In another embodiment of a nucleotide sequence of mouse TANGO 261, the nucleotide at position 106 is a guanine (G)(SEQ ID NO:19). In this embodiment, the amino acid at position 35 is a glutamate (E)(SEQ ID NO:20). In another embodiment of a nucleotide sequence of mouse TANGO 261, the nucleotide at position 106 is a cytosine

(C)(SEQ ID NO:76). In this embodiment, the amino acid at position 35 is aspartate (D)(SEQ ID NO:77). In another embodiment of a nucleotide sequence of mouse TANGO 261, the nucleotide at position 133 is a guanine (G)(SEQ ID NO:19). In this embodiment, the amino acid at position 44 is a glutamate (E)(SEQ ID NO:20). In another embodiment of a nucleotide sequence of mouse TANGO 261, the nucleotide at position 133 is a cytosine (C)(SEQ ID NO:78). In this embodiment, the amino acid at position 44 is aspartate (D)(SEQ ID NO:79).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the mouse TANGO 261 amino acid sequence in SEQ ID NO:20, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of mouse TANGO 261, nucleotides 5-652 (SEQ ID NO:112), encodes the mouse TANGO 261 amino acid sequence comprising amino acids 2-217 (SEQ ID NO:113).

In situ tissue screening was performed on mouse adult and embryonic tissue to analyze for the expression of mouse TANGO 261 mRNA. In the case of adult expression, a signal was observed in the cortex, olfactory bulb, caudate nucleus of the brain as well as in the brain stem. A weak signal was observed in the central grey matter of the spinal cord. A signal was observed in the ganglion cell layer of the eye and harderian gland. A signal was observed in the medulla of the adrenal gland. A moderate signal was observed in the cortex of the thymus. A signal was observed in the follicles of the spleen. A weak, ubiquitous signal was detected in the kidney, brown fat, and submandibular gland. A ubiquitous signal was detected in the liver, submandibular salivary gland, heart, colon, and in the cortical region of the adrenal gland. A signal was also observed in the labyrinth zone of the placenta and the mucosal epithelium of the bladder. A signal was also observed in the ovaries. No expression was observed in white fat, stomach, heart, lung, liver, lymph node, pancreas, skeletal muscle, testes, and small intestine.

In the case of embryonic expression, expression was seen in a number of tissues. At E13.5, a signal was observed in most tissues, the most noticeable exception being the liver which had a signal near background levels. The highest signal was observed in the ventricles of the brain. At E14.5, the strongest signal was observed in the eye. Weak to moderate signal was observed almost ubiquitously throughout the embryo. At E15.5 and E16.5, a strong signal was observed in the cortical region of the brain and the large vessels of the heart, descending aorta, and vessels associated with the umbilical cord. A moderate, ubiquitous signal was seen in the lung. A weak to moderate signal was observed in most other regions of the embryo. At E18.5, a very strong signal was observed in the eye,

specifically the developing retina. A strong signal was also seen in the large vessels of the heart, descending aorta, brown fat and submaxillary gland. A weak signal is observed in several other regions including the brain, intestinal tract, and the bladder. At P1.5, the signal had decreased to nearly background levels in most regions. The strongest signal was associated with the developing incisor teeth and the basio bone. A weak signal is also observed in the cortical and caudate regions of the brain.

Human and murine TANGO 261 sequences exhibit considerable similarity at the protein, nucleic acid, and open reading frame levels. An alignment (made using the ALIGN software (Myers and Miller (1989) CABIOS, ver. 2.0); BLOSUM 62 scoring matrix; gap penalties -12/-4), reveals a protein identity of 92.6%. The human and murine TANGO 261 full length cDNAs are 83.9% identical, as assessed using the same software and parameters as indicated (without the BLOSUM 62 scoring matrix). In the respective ORFs, calculated in the same fashion as the full length cDNAs, human and murine TANGO 261 are 87.4% identical.

Figure 8 depicts the alignment of the amino acid sequence of human TANGO 261 (SEQ ID NO:5) and a portion of murine TANGO 261 (SEQ ID NO:20). In this alignment, a (|) between the two sequences indicates an exact match.

Uses of TANGO 261 Nucleic Acids, Polypeptides, and Modulators Thereof

The TANGO 261 proteins and nucleic acid molecules of the invention have at least one "TANGO 261 activity" (also referred to herein as "TANGO 261 biological activity"). TANGO 261 activity refers to an activity exerted by a TANGO 261 protein or nucleic acid molecule on a TANGO 261 responsive cell *in vivo* or *in vitro*. Such TANGO 261 activities include at least one or more of the following activities: 1) interaction of a TANGO 261 protein with a TANGO 261-target molecule; 2) activation of a TANGO 261 target molecule; 3) modulation of cellular proliferation; 4) modulation of cellular differentiation; or 5) modulation of a signaling pathway. Thus, the TANGO 261 proteins, nucleic acids and/or modulators can be used for the treatment of a disorder characterized by aberrant TANGO 261 expression and/or an aberrant TANGO 261 activity, such as proliferative and/or differentiative disorders.

As TANGO 261 is expressed in the kidney, the TANGO 261 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues in which it is expressed. Such molecules can also be used to treat disorders associated with abnormal or aberrant metabolism or function of cells in the tissues in which it is expressed. Such can be used to

treat or modulate renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal disease, medullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atheroembolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

Because TANGO 261 is expressed in the reproductive tract, particularly in the ovaries, the TANGO 261 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues in which it is expressed. For example, the TANGO 261 polypeptides, nucleic acids and/or modulators thereof can be used modulate the function, morphology, proliferation and/or differentiation of the ovaries. For example. such molecules can be used to treat or modulate disorders associated with the ovaries, including, without limitation, ovarian tumors, McCune-Albright syndrome (polyostotic fibrous dysplasia). For example, the TANGO 261 polypeptides, nucleic acids and/or modulators can be used in the treatment of infertility.

As TANGO 261 exhibits expression in the lung, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat pulmonary (lung) disorders, such as atelectasis, pulmonary congestion or edema, chronic obstructive airway disease (e.g., emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis), diffuse interstitial diseases (e.g., sarcoidosis, pneumoconiosis, hypersensitivity pneumonitis, Goodpasture's syndrome, idiopathic pulmonary hemosiderosis, pulmonary alveolar proteinosis, desquamative interstitial pneumonitis, chronic interstitial pneumonia, fibrosing alveolitis, hamman-rich syndrome, pulmonary eosinophilia, diffuse interstitial fibrosis, Wegener's granulomatosis, lymphomatoid granulomatosis, and lipid pneumonia), or tumors (e.g., bronchogenic carcinoma, bronchiolovlveolar carcinoma, bronchial carcinoid, haematoma, and mesenchymal tumors).

As TANGO 261 exhibits expression in the spleen, TANGO 261 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of cells that form the spleen, e.g., cells of the splenic connective tissue, e.g., splenic smooth muscle cells and/or endothelial cells of the splenic blood vessels. TANGO 261 nucleic acids, proteins, and modulators thereof can also be used to modulate the proliferation, differentiation, and/or function of cells that are processed, *e.g.*, regenerated or phagocytized within the spleen, e.g., erythrocytes and/or B and T lymphocytes and macrophages. Thus, TANGO 261 nucleic acids, proteins, and modulators thereof can be used to treat spleen, *e.g.*, the fetal spleen, associated diseases and disorders. Examples of splenic diseases and disorders include *e.g.*, splenic lymphoma and/or splenomegaly, and/or phagocytotic disorders, *e.g.*, those inhibiting macrophage engulfment of bacteria and viruses in the bloodstream.

As TANGO 261 exhibits expression in the heart, TANGO 261 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, e.g., ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease.

As TANGO 261 exhibits expression in bone structures, TANGO 261 nucleic acids, proteins, and modulators thereof can be used to modulate the proliferation, differentiation, and/or function of bone and cartilage cells, *e.g.*, chondrocytes and osteoblasts, and to treat bone and/or cartilage associated diseases or disorders. Examples of bone and/or cartilage diseases and disorders include bone and/or cartilage injury due to for example, trauma (*e.g.*, bone breakage, cartilage tearing), degeneration (*e.g.*, osteoporosis), degeneration of joints, *e.g.*, arthritis, *e.g.*, osteoarthritis, and bone wearing.

In another example, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat disorders of the brain, such as cerebral edema, hydrocephalus, brain herniations, iatrogenic disease (due to, e.g., infection, toxins, or drugs), inflammations (e.g., bacterial and viral meningitis, encephalitis, and cerebral toxoplasmosis), cerebrovascular diseases (e.g., hypoxia, ischemia, and infarction, intracranial hemorrhage and vascular malformations, and hypertensive encephalopathy), and tumors (e.g., neuroglial tumors, neuronal tumors, tumors of pineal cells, meningeal tumors, primary and secondary lymphomas, intracranial tumors, and medulloblastoma). and to treat injury or trauma to the brain. Other examples of such brain and CNS related disorders include but are not limited to bacterial and viral meningitis, Alzheimers Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers (e.g., metastatic carcinoma of the brain, glioblastoma, lymphoma, astrocytoma, acoustic neuroma), hydrocephalus, and encephalitis.

In another example, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat hepatic (liver) disorders, such as jaundice, hepatic failure, hereditary hyperbiliruinemias (e.g., Gilbert's syndrome, Crigler-Naijar syndromes and Dubin-Johnson and Rotor's syndromes), hepatic circulatory disorders (e.g., hepatic vein thrombosis and portal vein obstruction and thrombosis) hepatitis (e.g., chronic active hepatitis, acute viral hepatitis, and toxic and drug-induced hepatitis) cirrhosis (e.g., alcoholic cirrhosis, biliary cirrhosis, and hemochromatosis), or malignant tumors (e.g., primary carcinoma, hepatoblastoma, and angiosarcoma).

In another example, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal diseasemedullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atheroembolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

In another example, as TANGO 261 exhibits expression in the brain, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat disorders of the brain, such as cerebral edema, hydrocephalus, brain herniations, iatrogenic disease (due to, e.g., infection, toxins, or drugs), inflammations (e.g., bacterial and viral meningitis, encephalitis, and cerebral toxoplasmosis), cerebrovascular diseases (e.g., hypoxia, ischemia, and infarction, intracranial hemorrhage and vascular malformations, and hypertensive encephalopathy), and tumors (e.g., neuroglial tumors, neuronal tumors, tumors of pineal cells, meningeal tumors, primary and secondary lymphomas, intracranial tumors, and medulloblastoma), and to treat injury or trauma to the brain. Other examples of such brain and CNS related disorders include, but are not limited to, bacterial and viral meningitis, Alzheimers Disease, cerebral toxoplasmosis, Parkinson's disease, multiple sclerosis, brain cancers (e.g., metastatic carcinoma of the brain, glioblastoma, lymphoma, astrocytoma, acoustic neuroma), hydrocephalus, and encephalitis.

In another example, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat prostate disorders, such as inflammatory diseases (*e.g.*, acute and chronic prostatitis and granulomatous prostatitis), hyperplasia (*e.g.*, benign prostatic hypertrophy or hyperplasia), or tumors (*e.g.*, carcinomas).

In another example, TANGO 261 polypeptides, nucleic acids, or modulators thereof, can be used to treat eye disorders, e.g., retinitis pigmentosa, cataract, retinalastoma, color blindness, conjunctivitis, myopia, dry eyes, keratoconus, glaucoma, macular degeneration, microphthalmia and anophthalmia, nystagmus, and trachoma.

<u>TANGO 262</u>

In another aspect, the present invention is based on the discovery of nucleic acid sequences which encode a novel family of proteins referred to herein as TANGO 262 proteins. Described herein are human TANGO 262 (SEQ ID NO:7), and murine TANGO 262 (SEQ ID NO:22) nucleic acid molecules and the corresponding polypeptides which the nucleic acid molecules encode (SEQ ID NO:8 and SEQ ID NO:23, respectively).

The TANGO 262 proteins and nucleic acid molecules comprise a family of molecules having certain conserved structural and functional features. As used herein, the term "family" is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Family members can be from either the same or different species. For example, a family can comprises two or more proteins of human origin, or can comprise one or more proteins of human origin and one or more of non-human origin. Members of the same family may also have common structural domains.

Also included within the scope of the present invention are TANGO 262 proteins having a signal sequence. As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 15 to 40 amino acid residues, preferably about 15-30 amino acid residues, and has at least about 60-80%, more preferably 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer.

In certain embodiments, a TANGO 262 family member has the amino acid sequence of SEQ ID NO:5, and the signal sequence is located at amino acids 1 to 19, 1 to

20, 1 to 21, 1 to 22 or 1 to 23. In such embodiments of the invention, the domains and the mature protein resulting from cleavage of such signal peptides are also included herein. For example, the cleavage of a signal sequence consisting of amino acids 1 to 21 of SEQ ID NO:8 (SEQ ID NO:32) results in a mature TANGO 262 protein corresponding to amino acids 22 to 226 of SEQ ID NO:8 (SEQ ID NO:33). The signal sequence is normally cleaved during processing of the mature protein.

In one embodiment, a TANGO 262 protein includes a signal sequence and is secreted.

Various features of human and mouse TANGO 262 are summarized below.

Human TANGO 262

Two clones were originally found in a fetal lung and kidney cell library, as ESTs with similarity to a C. elegans protein encoding gene. The full length sequence was eventually found in a stimulated kidney cell library. A cDNA clone, jthKa045g11, encoding full length human TANGO 262 was identified by screening a stimulated human kidney cell library by EST analysis. The 1682 nucleotide human TANGO 262 sequence (Figure 9A-9B; SEQ ID NO:7) includes an open reading frame which extends from nucleotide 322 to nucleotide 999 of SEQ ID NO:7 (SEQ ID NO:9) and encodes a 226 amino acid secreted protein depicted in Figure 9A-9B (SEQ ID NO:8).

In another embodiment, a cDNA encoding human TANGO 262 was identified by analyzing the sequences of clones present in a human a fetal lung library by EST analysis for sequences that encode wholly secreted or transmembrane proteins. This analysis led to the identification of a clone, jthKa045g11, comprising a 1510 nucleotide cDNA (SEQ ID NO:137). The open reading frame of this cDNA comprises nucleotides 325 to 1005 (SEQ ID NO:138), and encodes a transmembrane protein comprising the 226 amino acid sequence depicted in SEQ ID NO:8.

In one embodiment of a nucleotide sequence of human TANGO 262 the nucleotide at position 28 is a guanine (G)(SEQ ID NO:7). In this embodiment, the amino acid at position 2 is glutamate (E)(SEQ ID NO:8). In another embodiment of a nucleotide sequence of human TANGO 262, the nucleotide at position 28 is a cytosine (C)(SEQ ID NO:80). In this embodiment, the amino acid at position 2 is aspartate (D)(SEQ ID NO:81). In another embodiment of a nucleotide sequence of human TANGO 262, the nucleotide at position 483 is a guanine (G)(SEQ ID NO:7). In this embodiment. the amino acid at position 54 is a glutamate (E)(SEQ ID NO:8). In another embodiment of a nucleotide sequence of human TANGO 262, the nucleotide at position 483 is a cytosine (C)(SEQ ID

NO:82). In this embodiment, the amino acid at position 54 is aspartate (D)(SEQ ID NO:83). In another embodiment of a nucleotide sequence of human TANGO 262, the nucleotide at position 495 is a guanine (G)(SEQ ID NO:7). In this embodiment, the amino acid at position 58 is a glutamate (E)(SEQ ID NO:8). In another embodiment of a nucleotide sequence of human TANGO 262, the nucleotide at position 495 is a cytosine (C)(SEQ ID NO:84). In this embodiment, the amino acid at position 58 is aspartate (D)(SEQ ID NO:85).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the human TANGO 262 amino acid sequence in SEQ ID NO:8, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of human TANGO 262, nucleotides 325-999 (SEQ ID NO:114), encodes the human TANGO 262 amino acid sequence comprising amino acids 2-226 (SEQ ID NO:115).

Human TANGO 262 includes an signal sequence (amino acid 1 to about amino acid 21 of SEQ ID NO:8; SEQ ID NO:32) preceding the mature protein (about amino acid 22 to amino acid 226 of SEQ ID NO:8; SEQ ID NO:33). Human TANGO 262 protein, including the signal sequence, has a molecular weight of 24.6 kDa prior to post-translational modification. Mature human TANGO 262 protein has a molecular weight of 22.5 kDa after post-translational modification. The presence of a methionine residue at positions 53, 91, 111, 119, and 146 of SEQ ID NO:8 indicate that there can be alternative forms of human TANGO 262 of 174 amino acids of SEQ ID NO:8, 136 amino acids of SEQ ID NO:8, 116 amino acids of SEQ ID NO:8, 108 amino acids of SEQ ID NO:8, and 81 amino acids of SEQ ID NO:8, respectively.

In one embodiment, mouse TANGO 262 includes an extracellular domain at amino acids 22 to 226 of SEQ ID NO:22 (SEQ ID NO:).

A clone, EpT262, which encodes human TANGO 262 was deposited with the American Type Culture Collection (ATCC™, 10801 University Boulevard, Manassas, VA 20110-2209) on March 26, 1999, and assigned Accession Number 207176. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Figure 11 depicts a hydropathy plot of human TANGO 262. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. As shown in the hydropathy plot, the hydrophobic region of the plot which

corresponds to amino acid 1 to about amino acid 21 of SEQ ID NO:8 (SEQ ID NO:32) is the signal sequence of human TANGO 262. The cysteine residues (cys) are indicated by short vertical lines just below the hydropathy trace.

Northern analysis of human TANGO 262 mRNA expression revealed the presence of an approximately 1.8 kb transcript and an approximately 5.05 kb transcript that are expressed in a range of tissues including strong expression in heart; expression in the brain, skeletal muscle, kidney, liver, small intestine, lung, and placenta. No expression was detected in the colon, thymus, peripheral blood leukocytes, and spleen. The two transcripts likely represent alternative poly A site usage.

Human TANGO 262 is likely expressed in kidney, neuronal cells, placenta, bone, and fetal adrenal tissue, based on the origin of ESTs.

The human gene for TANGO 262 was mapped on radiation hybrid panels to the long arm of chromosome 14, in the region q23-q24. Flanking markers for this region are WI-6253 and WI-5815. The FNTB (fanesyltransferase) and MNAT1 (menage) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 12.

Mouse TANGO 262

A mouse homolog of human TANGO 262 was identified. A cDNA encoding mouse TANGO 262 was identified by analyzing the sequences of clones present in a mouse microglial cell cDNA library. This analysis led to the identification of a clone, jtmxa002h01, encoding mouse TANGO 262. The murine TANGO 262 cDNA of this clone is 1425 nucleotides long (Figure 10; SEQ ID NO:22). The open reading frame of this cDNA comprises nucleotides 89 to 766 of SEQ ID NO:22 (SEQ ID NO:24) and encodes the 226 amino acid mouse TANGO 262 secreted protein depicted in Figure 10 (SEQ ID NO:23).

In another embodiment, a mouse TANGO 262 clone includes comprises a 460 nucleotide cDNA (SEQ ID NO:139). The open reading frame of this cDNA comprises nucleotides 83 to 460 (SEQ ID NO:141), and encodes a transmembrane protein comprising the 126 amino acid sequence depicted in SEQ ID NO:140.

The signal peptide prediction program SIGNALP (Nielsen et al. (1997) *Protein Engineering* 10:1-6) predicted that mouse TANGO 262 includes a 21 amino acid signal peptide (amino acids 1 to about amino acid 21 of SEQ ID NO:23; SEQ ID NO:150) preceding the mature TANGO 262 protein (corresponding to about amino acid 22 to amino acid 226 of SEQ ID NO:23; SEQ ID NO:150). Mouse TANGO 262 protein, including the

signal sequence, has a molecular weight of 24.7 kDa prior to post-translational modification. Mature mouse TANGO 262 protein has a molecular weight of 22.5 kDa after post-translational modification. The presence of a methionine residue at positions 53, 91, 111, 113, 119, and 147 of SEQ ID NO:23 indicate that there can be alternative forms of mouse TANGO 262 of 174 amino acids of SEQ ID NO:23, 136 amino acids of SEQ ID NO:23, 116 amino acids of SEQ ID NO:23, 114 amino acids of SEQ ID NO:23, 108 amino acids of SEQ ID NO:23, and 80 amino acids of SEQ ID NO:23, respectively.

In one embodiment of a nucleotide sequence of mouse TANGO 262 the nucleotide at position 94 is a guanine (G)(SEQ ID NO:23). In this embodiment, the amino acid at position 2 is glutamate (E)(SEQ ID NO:24). In another embodiment of a nucleotide sequence of mouse TANGO 262, the nucleotide at position 94 is a cytosine (C)(SEQ ID NO:86). In this embodiment, the amino acid at position 2 is aspartate (D)(SEQ ID NO:87). In another embodiment of a nucleotide sequence of mouse TANGO 262, the nucleotide at position 250 is a guanine (G)(SEQ ID NO:23). In this embodiment, the amino acid at position 54 is a glutamate (E)(SEQ ID NO:24). In another embodiment of a nucleotide sequence of mouse TANGO 262, the nucleotide at position 250 is a cytosine (C)(SEQ ID NO:88). In this embodiment, the amino acid at position 54 is aspartate (D)(SEQ ID NO:89). In another embodiment of a nucleotide sequence of mouse TANGO 262, the nucleotide at position 262 is an adenine (A)(SEQ ID NO:23). In this embodiment, the amino acid at position 58 is a glutamate (E)(SEQ ID NO:24). In another embodiment of a nucleotide sequence of mouse TANGO 262, the nucleotide at position 262 is a cytosine (C)(SEQ ID NO:90). In this embodiment, the amino acid at position 58 is aspartate (D)(SEQ ID NO:91).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the mouse TANGO 262 amino acid sequence in SEQ ID NO:23, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of mouse TANGO 262, nucleotides 92-766 (SEQ ID NO:116), encodes the mouse TANGO 262 amino acid sequence comprising amino acids 2-226 (SEQ ID NO:117).

In situ tissue screening was performed on mouse adult and embryonic tissue to analyze for the expression of mouse TANGO 262 mRNA. Expression was widespread during the earlier embryonic ages examined. Expression in the limb, facial, and gut tissues suggested that skeletal muscle may be the predominant contributor to the signal observed in these areas. Strong expression was also seen in the brain and was localized to the area surrounding the lateral ventricles. Spinal cord and other regions of the brain had a

significant decrease or lack of expression. Mid and late stage embryos lacked the broad signal seen at earlier ages and had signal in a more defined pattern. The tissues lung, heart, kidney, eye, mucosal epithelium region of the stomach, and the intestinal tract all exhibited strong expression. The area of the brain in contact with the lateral ventricles remained high in expression until E18.5 and then became localized to the choroid plexus. Adult expression remained high in the gut with the stomach, small intestine, and colon all exhibiting strong expression. Kidney and adrenal gland also had expression, as did the choroid plexus as observed in the late stage embryos.

In the case of adult expression, the following results were obtained: A signal was observed in the brain in the choroid plexus of the lateral and 4th ventricles. A strong signal was observed in the mucosal epithelium of the stomach and the colon. A signal was observed in the region of the pericardium of the heart. A weak signal was observed in the ganglion layer of the eye and the harderian gland. A strong, ubiquitous signal was observed in the submandibular gland. A signal was observed in the cortical region of the kidney consistent with the pattern of glomeruli. There was also a ubiquitous signal in the medulla. A strong signal was observed in the cortical region of the adrenal gland. A strong signal was also obtained in the epithelium and villi of the small intestine. A signal was observed in the skeletal muscle/smooth muscle (particularly the diaphragm and peritoneum). A signal was observed in the mucosal epithelium and the serosa of the bladder. No expression was observed in the spinal cord, white fat, brown fat, lung, liver, thymus, lymph node, spleen, and pancreas.

In the case of embryonic expression, the following results were obtained: At E13.5, a signal was observed in a large number of tissues. The signal in the brain was very strong adjacent to the ventricles. The facial region, diaphragm, lung, kidney, and limbs exhibited a very strong signal. A broad expression signal pattern in the limbs suggested developing skeletal muscle. At E14.5, the signal was widely distributed throughout. Tissues lacking strong signal included the brain, except in the regions adjacent to ventricle, the spinal cord, and the liver. At E15.5, a strong signal was observed in the eye, lung, gut, kidney, and the digits of limbs. A signal was also seen in the whisker pads, brain adjacent to the ventricles, Meckel's cartilage, submaxillary gland, heart, and the peritoneum. At E16.5, the signal in the limbs and facial area had decreased to almost background levels suggesting a decrease or loss in signal from developing skcletal muscle. A strong signal was still observed in the eye, ventricle areas of the brain, whisker pads, Meckel's cartilage, submaxillary gland, heart, lung, and kidney. Signal was clearly observed in the mucosal portion of the stomach and the small intestine. At E18.5, the signal pattern is very similar to that observed at E16.5

with the noticeable exception being a significant decrease in signal in the brain adjacent to the ventricles and an increase in signal in the cortical and olfactory bulb areas. The continued decrease in possible muscle or connective tissue signal made the signal in the gut, small intestine and stomach, kidney, lung, and submaxillary gland even more pronounced. At P1.5, a strong signal was observed in the eye, submaxillary gland, kidney, the portion of the stomach containing the mucosal epithelium, and the intestinal tract. A less intense signal was seen in the upper and lower mandible, and the lung. The signal in the brain had decreased to almost background levels except in the choroid plexus.

Human and murine TANGO 262 sequences exhibit considerable similarity at the protein, nucleic acid, and open reading frame levels. An alignment (made using the ALIGN software (Myers and Miller (1989) CABIOS, ver. 2.0); BLOSUM 62 scoring matrix; gap penalties -12/-4), reveals a protein identity of 98.7% over the length of the mouse TANGO 226 protein (SEQ ID NO:23). The human and murine TANGO 262 full length cDNAs are 77.0% identical, as assessed using the same software and parameters as indicated (without the BLOSUM 62 scoring matrix). In the respective ORFs, calculated in the same fashion as the full length cDNAs, human and murine TANGO 262 are 88.5% identical.

Figure 12 depicts the alignment of the amino acid sequence of human TANGO 262 (SEQ ID NO:8) and the murine TANGO 262 amino acid sequence (SEQ ID NO:23). In this alignment, a (|) between the two sequences indicates an exact match.

Human TANGO 262 protein bears similarity *C. elegans* protein K10C3.4. Genbank Accession Number AC003687 appears to be the genomic sequence of human TANGO 262 (SEQ ID NO:).

Uses of TANGO 262 Nucleic Acids, Polypeptides, and Modulators Thereof

The TANGO 262 proteins and nucleic acid molecules of the invention have at least one "TANGO 262 activity" (also referred to herein as "TANGO 262 biological activity"). TANGO 262 activity refers to an activity exerted by a TANGO 262 protein or nucleic acid molecule on a TANGO 262 responsive cell *in vivo* or *in vitro*. Such TANGO 262 activities include at least one or more of the following activities: 1) interaction of a TANGO 262 protein with a TANGO 262-target molecule; 2) activation of a TANGO 262 target molecule; 3) modulation of cellular proliferation; 4) modulation of cellular differentiation; or 5) modulation of a signaling pathway. Thus, the TANGO 262 proteins, nucleic acids and/or modulators can be used for the treatment of a disorder characterized by aberrant TANGO 262 expression and/or an aberrant TANGO 262 activity, such as proliferative and/or differentiative disorders.

TANGO 262 proteins, nucleic acids and/or modulators of the invention are useful in the treatment of disorders of the kidney, nervous system, bone, and adrenal gland.

As TANGO 262 is expressed in the kidney, the TANGO 262 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues in which it is expressed. Such molecules can also be used to treat disorders associated with abnormal or aberrant metabolism or function of cells in the tissues in which it is expressed. Such can be used to treat or modulate renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal disease, medullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atheroembolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

As TANGO 262 exhibits expression in the lung, TANGO 262 polypeptides, nucleic acids, or modulators thereof, can be used to treat pulmonary (lung) disorders, such as atelectasis, pulmonary congestion or edema, chronic obstructive airway disease (e.g., emphysema, chronic bronchitis, bronchial asthma, and bronchiectasis), diffuse interstitial diseases (e.g., sarcoidosis, pneumoconiosis, hypersensitivity pneumonitis, Goodpasture's syndrome, idiopathic pulmonary hemosiderosis, pulmonary alveolar proteinosis, desquamative interstitial pneumonitis, chronic interstitial pneumonia, fibrosing alveolitis, hamman-rich syndrome, pulmonary eosinophilia, diffuse interstitial fibrosis, Wegener's granulomatosis, lymphomatoid granulomatosis, and lipid pneumonia), or tumors (e.g., bronchogenic carcinoma, bronchiolovlveolar carcinoma, bronchial carcinoid, hamartoma, and mesenchymal tumors).

As TANGO 262 exhibits expression in the heart, TANGO 262 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, e.g., ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease.

As TANGO 262 exhibits expression in the small intestine, TANGO 262 polypeptides, nucleic acids, or modulators thereof, can be used to treat intestinal disorders, such as ischemic bowel disease, infective enterocolitis, Crohn's disease, benign tumors, malignant tumors (e.g., argentaffinomas, lymphomas, adenocarcinomas, and sarcomas), malabsorption syndromes (e.g., celiac disease, tropical sprue, Whipple's disease, and abetalipoproteinemia), obstructive lesions, hernias, intestinal adhesions, intussusception, or volvulus.

In another example, TANGO 262 polypeptides, nucleic acids, or modulators thercof, can be used to treat hepatic (liver) disorders, such as jaundice, hepatic failure, hereditary hyperbiliruinemias (e.g., Gilbert's syndrome, Crigler-Naijar syndromes and Dubin-Johnson and Rotor's syndromes), hepatic circulatory disorders (e.g., hepatic vein thrombosis and portal vein obstruction and thrombosis) hepatitis (e.g., chronic active hepatitis, acute viral hepatitis, and toxic and drug-induced hepatitis) cirrhosis (e.g., alcoholic cirrhosis, biliary cirrhosis, and hemochromatosis), or malignant tumors (e.g., primary carcinoma, hepatoblastoma, and angiosarcoma).

In another example, TANGO 262 polypeptides, nucleic acids, or modulators thereof, can be used to treat renal (kidney) disorders, such as glomerular diseases (e.g., acute and chronic glomerulonephritis, rapidly progressive glomerulonephritis, nephrotic syndrome, focal proliferative glomerulonephritis, glomerular lesions associated with systemic disease, such as systemic lupus erythematosus, Goodpasture's syndrome, multiple myeloma, diabetes, neoplasia, sickle cell disease, and chronic inflammatory diseases), tubular diseases (e.g., acute tubular necrosis and acute renal failure, polycystic renal diseasemedullary sponge kidney, medullary cystic disease, nephrogenic diabetes, and renal tubular acidosis), tubulointerstitial diseases (e.g., pyelonephritis, drug and toxin induced tubulointerstitial nephritis, hypercalcemic nephropathy, and hypokalemic nephropathy) acute and rapidly progressive renal failure, chronic renal failure, nephrolithiasis, vascular diseases (e.g., hypertension and nephrosclerosis, microangiopathic hemolytic anemia, atherocmbolic renal disease, diffuse cortical necrosis, and renal infarcts), or tumors (e.g., renal cell carcinoma and nephroblastoma).

TANGO 266

In another aspect, the present invention is based on the discovery of nucleic acid sequences which encode a novel family of proteins referred to herein as TANGO 266 proteins. Described herein is a human TANGO 266 (SEQ ID NO:10) nucleic acid molecule and the corresponding protein which the nucleic acid molecule encodes (SEQ ID NO:11).

TANGO 266 proteins and nucleic acid molecules comprise a family of molecules having certain conserved structural and functional features. As used herein, the term "family" is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Family members can be from either the same or different species. For example, a family can comprises two or more proteins of human origin, or can comprise one or more proteins of human origin and one or more of non-human origin. Members of the same family may also have common structural domains.

Also included within the scope of the present invention are TANGO 266 proteins having a signal sequence. As used herein, a "signal sequence" includes a peptide of at least about 15 or 20 amino acid residues in length which occurs at the N-terminus of secretory and membrane-bound proteins and which contains at least about 70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a signal sequence contains at least about 15 to 40 amino acid residues, preferably about 15-30 amino acid residues, and has at least about 60-80%, more preferably 65-75%, and more preferably at least about 70% hydrophobic residues. A signal sequence serves to direct a protein containing such a sequence to a lipid bilayer.

In certain embodiments, a TANGO 266 family member has the amino acid sequence of SEQ ID NO:11, and the signal sequence is located at amino acids 1 to 17, 1 to 18, 1 to 19, 1 to 20 or 1 to 21. In such embodiments of the invention, the domains and the mature protein resulting from cleavage of such signal peptides are also included herein. For example, the cleavage of a signal sequence consisting of amino acids 1 to 19 of SEQ ID NO:11 (SEQ ID NO:34) results in a mature TANGO 266 protein corresponding to amino acids 20 to 105 of SEQ ID NO:11 (SEQ ID NO:35). The signal sequence is normally cleaved during processing of the mature protein.

Thus, in one embodiment, a TANGO 266 protein includes a signal sequence and is secreted.

HUMAN TANGO 266

A sequence encoding human TANGO 266 was identified by screening a human adrenal gland library by EST analysis. The 1422 nucleotide human TANGO 266 sequence (Figure 14; SEQ ID NO:10) includes an open reading frame which extends from nucleotide 49 to nucleotide 363 of SEQ ID NO:10 (SEQ ID NO:12) and encodes a 105 amino acid secreted protein depicted in Figure 14 (SEQ ID NO:11).

In another embodiment, a human TANGO 266 clone includes comprises a 422 nucleotide cDNA (SEQ ID NO:142). The open reading frame of this cDNA comprises nucleotides 56 to 373 (SEQ ID NO:243), and encodes a transmembrane protein comprising the 105 amino acid sequence depicted in SEQ ID NO:11.

In one embodiment of a nucleotide sequence of human TANGO 266 the nucleotide at position 129 is a guanine (G)(SEQ ID NO:10). In this embodiment, the amino acid at position 27 is glutamate (E)(SEQ ID NO:11). In another embodiment of a nucleotide sequence of human TANGO 266, the nucleotide at position 129 is a cytosine (C)(SEQ ID NO:92). In this embodiment, the amino acid at position 27 is aspartate (D)(SEQ ID NO:93) In another embodiment of a nucleotide sequence of human TANGO 266, the nucleotide at position 216 is an adenine (A)(SEQ ID NO:10). In this embodiment, the amino acid at position 56 is a glutamate (E)(SEQ ID NO:11). In another embodiment of a nucleotide sequence of human TANGO 266, the nucleotide at position 216 is a cytosine (C)(SEQ ID NO:94). In this embodiment, the amino acid at position 56 is aspartate (D)(SEQ ID NO:95). In another embodiment of a nucleotide sequence of human TANGO 266, the nucleotide at position 222 is a guanine (G)(SEQ ID NO:10). In this embodiment, the amino acid at position 58 is a glutamate (E)(SEQ ID NO:11). In another embodiment of a nucleotide sequence of human TANGO 266, the nucleotide at position 222 is a cytosine (C)(SEQ ID NO:96). In this embodiment, the amino acid at position 58 is aspartate (D)(SEQ ID NO:97).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the human TANGO 266 amino acid sequence in SEQ ID NO:11, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of human TANGO 266, nucleotides 52-363 (SEQ ID NO:18), encodes the human TANGO 216 amino acid sequence comprising amino acids 2-105 (SEQ ID NO:119).

Human TANGO 266 includes a signal sequence (amino acid 1 to about amino acid 19 of SEQ ID NO:11; SEQ ID NO:34) preceding the mature protein (about amino acid 20 to amino acid 105 of SEQ ID NO:11; SEQ ID NO:35). Human TANGO 266 protein, including the signal sequence, has a molecular weight of 11.7 kDa prior to post-translational modification. Mature human TANGO 266 protein has a molecular weight of 9.7 kDa after post-translational modification. The presence of a methionine residue at positions 10, 49, and 98 of SEQ ID NO:11 indicate that there can be alternative forms of human TANGO 266 of 96 amino acids, 57 amino acids, and 8 amino acids, respectively.

A clone, EpT266, which encodes human TANGO 266 was deposited with the American Type Culture Collection (ATCC™, 10801 University Boulevard, Manassas, VA 20110-2209) on March 26, 1999, and assigned Accession Number 207176. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

Figure 11 depicts a hydropathy plot of human TANGO 266. Relatively hydrophobic residues are above the horizontal line, and relatively hydrophilic residues are below the horizontal line. As shown in the hydropathy plot, the hydrophobic region of the plot which corresponds to amino acid 1 to about amino acid 19 of SEQ ID NO:11 (SEQ ID NO:34) is the signal sequence of human TANGO 266. The cysteine residues (cys) are indicated by short vertical lines just below the hydropathy trace.

Northern analysis of human TANGO 266 mRNA expression revealed the presence of an approximately 1.7 kb transcript that is expressed in a range of tissues including very strong expression in placenta; and weak expression in heart. An additional Northern was performed on human TANGO 266 in which strong expression was detected in the adrenal medulla and testis, and moderate expression was detected in the adrenal cortex. No expression was detected in the brain, lung, liver, skeletal muscle, kidney, and pancreas.

In situ tissue screening was performed on mouse adult and embryonic tissue to analyze for the expression of human TANGO 266 mRNA. Consistent with the Northern results obtained above, expression was seen in the ovarian stroma and placenta. The pattern of the signal suggested expression by a component of the vasculature. A stronger signal was observed in the testes. The pattern was multifocal and did not suggest expression by seminiferous tubules. Photoemulsion can be used to determine the exact cellular component of these tissues expressing human TANGO 266 mRNA.

Specifically, in the case of adult expression, a strong, multifocal signal was detected in the testes. A moderate signal was detected in the placenta. No expression was detected in the following tissues: brain (cerebellum), submandibular gland, heart, liver, kidney, colon, small intestine, and spleen.

Example 1: Isolation And Characterization of Human TANGO 266 cDNAs

A human TANGO 266 cDNA was isolated from a human adrenal gland cDNA library. A cDNA library from human adult adrenal gland RNA was constructed and sequenced by automated high throughput single pass sequencing, and individual clones

analyzed for homology to known proteins. A cDNA clone (TANGO 266) was found initially to have significant homology only to venom protein A (VPRA), found in high abundance in the venom of the black mamba (Dendroaspis polylepsis)(Schweitz, H., Didard, J. & Lazdunski, M. (1990) *Toxicon* 28, 847-856)(Boisbouvier, J. *et al.* (1998) *J. Mol. Biol.* 283, 205-219). TANGO 266 was found to be 58% identical to VPRA over the 81 residues of reported amino acid sequence (Figure 19A). Recently, a similar protein (Bv8) was isolated from skin secretions of the frog *Bombina Variegata* (Mollay, C. *et al.* (1999) *Eur. J. Pharmacol.* 374, 189-196) and the peptide sequence was used to clone the frog, mouse and human Bv8 cDNAs (Wechselberger, C. *et al.* (1999) *FEBS Lett.* 462, 177-181). The partial human Bv8 sequence reported was compared to that of TANGO 266 and found have 45% identity over the length of the published sequence.

Human TANGO 266 protein bears similarity to *Dendroaspis polypepis polypepis* venom protein A (SwissProt Accession Number P25687; Joubert and Strydom (1980) *Hoppe Seylers Z Physiol. Chem.* 361:1787-94). Figure 16 depicts the alignment of the amino acid sequence of human TANGO 266 (SEQ ID NO:11) and *Dendroaspis polypepis polypepis* venom protein A (SEQ ID NO:36). In this alignment, a (•) between the two sequences indicates an exact match. The cysteines at residues at positions 26, 32, 38, 50, 60, 78, 80, 86, and 96 of human TANGO 266 (SEQ ID NO:11) are conserved between human TANGO 266 and *Dendroaspis polypepis polypepis* venom protein A (SEQ ID NO:36), suggesting that these cysteines form disulfide bonds. A cysteine at amino acid position 37 in TANGO 266 is not found at the corresponding position in *Dendroaspis polypepis polypepis* venom protein A. However, a tenth cysteine occurs four residues beyond the corresponding position. This tenth cysteine residue is likely able to interact with its partner from either position.

Comparison of mouse Bv8 variant 3 to VPRA and TANGO 266 is shown in Fig 19A. Mouse Bv8 is closer in homology to VPRA than TANGO 266 (Figure 19B), with 60% identity over the region of the VPRA peptide sequence, whereas TANGO 266 shares 54% identity with VPRA. The primary structure of TANGO 266 is similar to Bv8 and VPRA, with identical amino terminal sequences (AVITGAC) and conservation of 10 cysteines in the mature protein, with the exception of VPRA, which lacks the first cysteine. The complete TANGO 266 cDNA (1,422 bp) encodes a 105 residue protein with a predicted molecular mass of 11,714 Daltons.

Example 2: Determination of TANGO 266 as Secreted Protein

To determine if the signal peptide prediction correctly determined that TANGO 266 is a secreted protein, cell lines were transfected with TANGO 266 cDNA and subjected to a secretion assay, and their supernatants were probed with rabbit anti human TANGO 266 peptide polyclonal antisera (as discussed below). 293 cells were transfected with expression vectors carrying TANGO 266 Fc-tagged fusion protein, alkaline phosphatase (AP) tagged fusion protein, or with a retroviral vector expressing the native protein. Media from transfected cells was collected and evaluated by Western for presence of secreted protein (Figure 19C). In all instances polyclonal anti-TANGO 266 recognized native or tagged protein. In addition, TANGO 266 could be detected in media of 3T3 cells infected with a retrovirus expressing native TANGO 266, but not in control cells infected with an empty vector. The procedures utilized for creation of fusion proteins, for production of the anti-TANGO 266 antibody, and for testing protein secretion, are as follows:

Creation of TANGO 266 Fusion Proteins

TANGO 266 was amplified by PCR and cloned into expression vectors containing different epitope tags. The following oligos were used:

P1: 5' TTTTTGAATTCACCGCCATGAGAGGTGCCACGCGAG 3' (SEQ ID NO: 151)
P2: 5' TTTTTCTCGAGAAAATTGATGTTCTTCAAGTCCA 3' (SEQ ID NO:152)
P3: 5' TTTTTAGATCTGCTGTGATCACAGGGGCC 3' (SEQ ID NO:153)
P4: 5' TTTTTCTCGAGCTAAAAATTGATGTTCTTCAAGTC 3' (SEQ ID NO:154)

TANGO 266 was amplified with P1 (contains EcoRI site and Kozak sequence) and P2 (contains XhoI site) and cloned in frame into the EcoRI and XhoI sites of the pMEAP3 vector 5' of alkaline phosphatase (TANGO 266-AP). Using the same sites TANGO 266 was also cloned into pcDNA3.1 containing either the sequence encoding for the Fc part of hIgG1 or a FLAG epitope adding the Fc (TANGO 266-Fc) or Flag (TANGO 266-Flag) sequence in frame to the 3' end of TANGO 266. Oligos P3 and P4 were used to clone TANGO 266 (without signal peptide) into the Bgl II and XhoI cloning sites of plasmid APTag3, 3' of alkaline phosphatase and in frame (AP-TANGO 266).

Production of Anti-TANGO 266 Antibody

Polyclonal anti-TANGO 266 was produced in rabbits using the peptide PLGREGEECHPGSHK (SEQ ID NO:). Antibody was peptide affinity purified from 12 week bleeds.

Protein Secretion Assay

The sequenced DNA constructs were transiently transfected into HEK 293T cells in 150mM plates using Lipofectamine (GIBCO/BRL) according to the manufacturer's protocol. 72 hours post-transfection, the serum-free conditioned media (OptiMEM, Gibco/BRL) were harvested, spun and filtered. Alkaline phosphatase activity in conditioned media was quantitated using an enzymatic assay kit (Phospalight) according to the manufacturer's instructions. Conditioned medium samples were analyzed by SDS-PAGE followed by Western blot using polyclonal anti-peptide antibodies to TANGO 266 as described previously. Isolation of the TANGO 266-Fc was performed with a one step purification scheme utilizing the affinity of the human IgG1 Fc domain to Protein A. The conditioned media was passed over a POROS A column (4.6 X 100 mm, PerSeptive Biosystems); the column was then washed with PBS, pH 7.4 and eluted with 200 mM glycine, pH 3.0. Samples were dialyzed against PBS, pH 7.4 at 4°C with constant stirring. The buffered exchanged material was then sterile filtered (0.2 micrometers, Millipore) and frozen at -80°C.

Example 3: TANGO 266 Tissue Distribution

Total RNA was prepared from various human tissues by a single step extraction method using RNA STAT-60 according to the manufacturer's instructions (TelTest, Inc). Each RNA preparation was treated with DNase I (Ambion) at 37°C for 1 hour. DNAse I treatment was determined to be complete if the sample required at least 38 PCR amplification cycles to reach a threshold level of flourescence using β-2 microglobulin as an internal amplicon reference. The integrity of the RNA samples following DNase I treatment was confirmed by agarose gel electrophoresis and ethidium bromide staining. After phenol extraction cDNA was prepared from the sample using the SuperScript™ Choice System following the manufacturer's instructions (GibcoBRL). A negative control of RNA without reverse transcriptase was mock reverse transcribed for each RNA sample. Expression was measured by TaqMan® quantitative PCR (Perkin Elmer Applied Biosystems) in cDNA prepared from the following normal human tissues: cecum, colon ascending, colon descending, colon transverse, duodenum, esophagus, ileocecum, ileum, jejunum, liver, rectum, stomach, heart, kidney, liver, pancreas, placenta, skeletal muscle, ovary, prostate, small intestine, testis, and adrenal tissue.

Each TANGO 266 gene probe was labeled using FAM (6-carboxyfluorescein), and the β2-microglobulin reference probe was labeled with a different fluorescent dye, VIC

(forward and reverse primers, and TaqMan probe, were designed by PrimerExpress software (PE Biosystems) based on the sequence of each gene). The differential labeling of the target gene and internal reference gene thus enabled measurement in the same well. Forward and reverse primers and the probes for both β2-microglobulin and target gene were added to the TaqMan® Universal PCR Master Mix (PE Applied Biosystems). Although the final concentration of primer and probe could vary, each was internally consistent within a given experiment. A typical experiment contained 200nM of forward and reverse primers plus 100nM probe for β-2 microglobulin and 600 nM forward and reverse primers plus 200 nM probe for the target gene. TaqMan matrix experiments were carried out on an ABI PRISM 7700 Sequence Detection System (PE Applied Biosystems).

The following method was used to quantitatively calculate gene expression: The threshold cycle (Ct) value was defined as the cycle at which a statistically significant increase in flourescence was detected. A lower Ct value was indicative of a higher mRNA concentration. The Ct value of the kinase gene was normalized by subtracting the Ct value of the β-2 microglobulin gene to obtain a ?Ct value using the following formula: $_{\Delta}Ct=Ct_{kinase} - Ct_{\beta\text{-}2\ microglobulin}$. Expression was then calibrated against a cDNA sample showing a comparatively low level of expression of the kinase gene. The $_{\Delta}$Ct value for the calibrator sample was then subtracted from $_{\Delta}$Ct for each tissue sample according to the following formula: $_{\Delta}Ct={}_{\Delta}Ct_{-sample} - {}_{\Delta}Ct\text{-calibrator}$. Relative expression was then calculated using the arithmetic formula given by $2^{-\Delta\Delta_{Ct}}$.

TANGO 266 gene expression was as follows: No expression was detected in colon ascending, colon descending, colon transverse, duodenum, esophagus, ileocecum, ileum, jejunum, liver, rectum, stomach, kidney, liver, and pancreas. Trace levels of expression were detected in small intestine, which shall serve as the baseline level of expression, relative to which other levels are compared. Skeletal muscle, heart, and prostate reveal levels of expression about five times greater than the level of expression in small intestine. Cecum, placenta, and adrenal tissue reveal levels of expression about 40-50 times greater than the level of expression in small intestine. Testis revealed a level of expression about 250 times stronger than the level of expression in small intestine, and in ovary the expression was about 500 times stronger than the level of expression in small intestine.

Example 4: Screening of Mouse Tissues for TANGO 266 Binding Sites

To identify potential sites of action of TANGO 266, mouse tissues sections were screened for binding sites using TANGO 266 alkaline phosphatase fusion proteins. Alkaline phosphatase was fused in frame either to the N-terminus (AP-TANGO 266) or the C-

terminus (TANGO 266-AP) of TANGO 266. Binding of TANGO 266-AP (as well as AP-TANGO 266) to scattered cells in bone marrow and in the red pulp of spleen was detected. Alkaline Phosphatase (AP) by itself was used as control and did not bind to spleen and bone marrow. The morphology of cells bound by TANGO 266 was reminiscent of cells of the monocyte/macrophage lineage and prompted an analysis of the binding of TANGO 266 to isolated bone marrow derived macrophages. TANGO 266-AP, but not AP by itself, bound to macrophages cultured *in vitro* for 3 days in the presence of M-CSF (macrophage colony stimulating factor). The binding studies were performed as follows. The isolation of bone marrow derived macrophages is also described below:

Binding studies using alkaline phosphatase fusion proteins were done as described in Cheng and Flanagan, *Cell* 79:157-168. Briefly, 8 μM cyrostat sections were prepared from tissues embedded in OCT and frozen in liquid nitrogen. Sections were thawed, washed once in HBHA (Hank's balanced salt solution supplemented with 20 mM Hepes, pH 7, 0.05% BSA and 0.1% sodium azide) and incubated with alkaline phospatase fusion proteins for one hour in a humidified chamber. Sections were washed 6 times in HBHA, fixed in acetone/paraformaldehyde, washed 3 x in HBS (20 mM Hepes, pH 7.5, 150 mM NaCl) and developed using BCIP/NBT substrate solution (100 mM Tris-HCl, pH 9.5, 100 mM NaCl, 5 mM MgCl, 0.17 mg/ml BCIP and 0.33 mg/ml NBT).

Bone marrow derived macrophages were obtained by culturing nucleated bone marrow cells (see the following section) with 50 ng/ml M-CSF on cover slips in 6-well plates. After three days, non-adherent cells were removed and adherent cells on cover slips were fixed in acetone and air-dried.

Example 5: Analysis of the Effect of TANGO 266 on Mononuclear Bone Marrow Cells

The results of the binding studies also prompted an analysis of the effect of purified TANGO 266-Fc on mononuclear bone marrow cells. Cells were cultured in the presence of TANGO 266-Fc for three days and mitogenic activity was measured by ^{3}H thymidine incorporation. TANGO 266-Fc was shown to induce a concentration-dependent increase in the mitogenic response. Maximal ^{3}H thymidine incorporation was detected at about 1500 ng/ml. A control-Fc fusion protein had no effect on the mitogenic response making it unlikely that the Fc part of the protein is responsible for the observed effect. Moreover, heat inactivation of TANGO 266-Fc (10 min at 95 degrees Celsius) abolished the mitogenic response ruling out the possibility that the functional response elicited by TANGO 266-Fc is due to endotoxin contamination in the protein preparation.

Culturing of mononuclear bone marrow cells (described below) in the presence of TANGO 266-Fc not only resulted in a mitogenic response but also in morphological changes. Large numbers of adherent cells of macrophage-like morphology were observed in cultures treated with 266-Fc but only few if any adherent cells were detected in cultures treated with culture medium only, control-Fc or heat-inactivated TANGO 266-Fc. Immuno-fluorescence analysis (discussed briefly below) showed that the adherent cell population was positive for Mac-1, a marker specific for the myeloid lineage and F4/80, a marker specific for macrophages indicating that the adherent cells are macrophages. This was further confirmed by FACS analysis using a range of different lineage markers. The adherent cell population stimulated by TANGO 266-Fc is Mac1+, F4/80+, Gr-1 low, B220- and CD3-. In summary, the above data show that TANGO 266-Fc stimulates a mitogenic response in mononuclear bone marrow cells, and the proliferation and differentiation of macrophages.

Culturing Bone Marrow Cells

Bone marrow was harvested from femurs of 4 to 6 week old C57BL6 mice and passed over a mouse density centrifugation medium (LympolyteM, Cedarlane laboratories, Ontario) to isolate nucleated cells. For the 3H thymidine incorporation assay, 0.5 to 1 x 105 nucleated cells were incubated in a total volume of 0.2 ml in individual of 96-well plates containing dilutions of TANGO 266 for 72 h. The culture medium used was McCoy's 5A mdium supplemented with 15 % fetal calf serum and antibiotics. During the last 6 hours of culture, cells were pulse labeled with 0.5 μCi 3H thymidine (5 Ci/mmol sp. act.) and 3H thymidine incorporation was quantified by scintillation counting as described.

Flow cytometry and Immuno-fluorescence

For flow cytometry analysis cultures were set up in 6-well plates. Adherent cells were detached in Versene, washed and then incubated for 60 min with 10 μg/ml of the FITC-conjugated marker antibodies. Cells were then washed and analyzed with a FACSCaliber flow cytometer. For in situ fluorescence analysis adherent cells grown on chamberslides were fixed in acetone, washed in PBS and incubated for 60 minutes with FITC-conjugated marker antibodies in a humidified staining chamber. Slides were washed in PBS, mounted with cover slips and analyzed under a fluorescence microscope.

Example 6: *In vivo* TANGO 266 Expression

To study the consequences of TANGO 266 expression *in vivo* (described below), we overexpressed TANGO 266 in the hematopoietic system of mice. To this end, hematopoietic progenitor cells from SJL mice were transduced with a retroviral vector carrying TANGO 266 (MSP-TANGO 266) or an empty control vector pMSCVpac (MSP). Transduced cells were then transplanted into sublethally irradiated C57Bl6 mice and allowed to reconstitute the hematopietic system. Two months after transplant, animals were sacrificed. Blood, bone marrow and spleen were analyzed by flow cytometry with different hematopoietic lineage markers including B220, IgD, CD3, NK1.1, Mac1, Gr-1 and F4/80. CD45.1, a marker specific for donor derived cells, was used as an indicator for the reconstitution efficiency.

The reconstitution efficiency was similar for all animals (about 90%). No differences in the distribution of the hematopoietic lineages were seen in blood and bone marrow between mice reconstituted with MSP-TANGO 266 transduced bone marrow (MSP-TANGO 266 mice) versus mice reconstituted with MSP transduced bone marrow (MSP mice). However, whereas the distribution of B220+, CD3+, NK1.1+ and Gr-1 positive cells was similar in the spleen of MSP-TANGO 266 mice and MSP mice, a higher percentage of Mac1/F4/80 double positive cells was observed in the spleen of MSP-TANGO 266 mice. This Mac1/F4/80 double positive population was hardly detectable in MSP control animals but was clearly visible in MSP-TANGO 266 animals. Mac1 expression was higher on this population compared to the F4/80 negative population. These results indicate that overexpression of TANGO 266 in the hematopoietic system in mice results in an increase of macrophages in the spleen.

In vivo animal studies

The full length human Tango 266 cDNA was cloned into pMSCVpac (MSP), a virus containing a PGK promoter driven the puromycin resistance gene. Control virus was the empty virus. The viruses were produced in the 293-EBNA cells by transfecting the retroviral plasmid with two PN8ε vectors, one containing the gag/pol construct, PN8ε gagpol, from the murine moloney leukemia virus (MMLV) and the other the VSV-G envelop, PN8ε VSV-G. Viral supernatants were collected 48 hours, 72 hours and 96 hours after transfection, filtered and centrifuged at 4C at 50,000 x g (25,000 rpm) for 2 hr. Concentrated virus pellets were resuspended in culture medium, shaken and frozen at -80C until transduction.

Donor mouse bone marrow cells were collected 4 days after treatment with 5-fluorouracil (5-FU), immunopurified for CD3e, CD11b, CD45R and Ly-6G negative cells, prestimulated for two days, infected for one day with the viral supernatant in the presence of recombinant mouse interleukin-3, recombinant mouse interleukin-6 (rmIL6), recombinant mouse stem cell factor (rmSCF), recombinant mouse fms-like tyrosine kinase-3 ligand (rmFlt-3L) and mouse thrombopoietin (mTPO) and then collected and injected into lethally irradiated recipient mice.

Example 8: Analysis of Progenitor Cells to Determine TANGO 266 Effect

In recent years culture conditions have been developed that allow human bone marrow CD34+ progenitors to expand *in vitro* and to differentiate into antigen presenting cells. (Zandstra, P.W., *et al* (1997). *Proc. Natl. Acad. Sci.* USA 94, 4698-4703; Bhatia, M. *et al.* (1997) *J. Exp. Med.* 186, 619-624; and Banchereau, J., & Steinman, R.M. (1998) Nature 392, 245-252.) CD34+ human bone marrow cells were cultured in serum free media in the presence of Flt-3 ligand, SCF, IL-3 and IL-6 in the presence or absence of TANGO 266-Fc. Under these conditions, total cell numbers in cytokines alone or with a control Fc fusion protein increased 200 – 400 fold. TANGO 266-Fc increased the proportion of adherent cells in expanded human bone marrow CD34+ cell cultures in a dose dependent manner. The morphology of the adherent cells was suggestive of cells differentiating into the monocyte / macrophage lineage.

Cells were assessed for stage of differentiation using CD34, an early hematopoietic progenitor marker, and CD14 and CD16 which are expressed by cells that have differentiated into the monocyte/macrophage lineage. CD14 is a functional receptor on cells of the monocytic lineage for bacterial lipopolysaccharide, and for clearance and phagocytosis of apoptotic cells. The addition of TANGO 266-Fc increased the number of cells expressing CD16. The addition of TANGO 266-Fc greatly decreased the percentage of CD34+ / CD14- cells, and increased CD34-/CD14+ cells after 14 days of culture, suggesting that TANGO 266 acts on early progenitors to induce differentiation into the monocyte lineage. This affect was not evident in media alone, with a control Fc fusion protein, or with heat inactivated TANGO 266-Fc. Total cell number after 2 weeks in culture increased 1.5 – 2.2 fold compared to media alone or in presence of a control Fc protein. The total number of CD34+ cells in culture dropped 10 fold, with a concomitant 3 fold increase in the number of CD14+ cells when cultured in the presence of 200 ng ml^{-1} TANGO 266-Fc compared to a control Fc. This effect was seen in a dose dependent manner

in a range of 1 – 500 ng ml^{-1} when cultured for a 2 week period. The human bone marrow cell culture and analysis is described as follows:

Human bone marrow CD34+ cell culture and analysis

Adult human bone marrow cells selected for expression of CD34 were purchased from Purecell (Foster City, CA). Cells (4 X 10^3 ml^{-1}) were cultured for 14 days in serum free media containing cytokines (StemCell Tech., Vancouver, B.C., Canada) Flt-3 ligand (100 ng ml^{-1}), SCF (100 ng ml^{-1}), IL-3 (10 ng ml^{-1}) and IL-6 (10 ng ml^{-1}) in a humidified 5%CO_2 incubator at 37°C. Non adherent cells were collected and adherent cells removed by with a cell lifter after incubation in Versene (Gibco/BRL, Grand Island, NY), washed and blocked with 1 mg ml^{-1} human gamma globulin (Gamimune; Miles Inc, Elkhart, IN). Total viable cell count was determined by trypan blue exclusion. Fluorescein isothiocyanate (FITC) labeled anti-CD14 and anti-CD16, and phycoerythrin (PE) labeled anti-CD34 were obtained from Pharmingen. After dilution in PBS cells were analyzed by FACSCaliber flow cytometer (Becton Dickinson, Franklin Lakes, NJ).

Example 9: Mapping Results of TANGO 266

The TANGO 266 nucleic acid sequence bears homology to a marker called SHGC-16135, which is known to map to 1p21. 1p21 is a locus for a disorder known as osteopetrosis, autosomal dominant, type II, the mapping of which was discovered during a study of an extended family with type II disorder (Van Hul, W. *et al* (1997) *Medizinische Genetik* 9: 8). In the study, linkage between the disorder and to microsatellite markers in the 1p21 region was demonstrated. The chromosomal region was further analyzed, within which was discovered the gene for macrophage colony stimulating factor (CSF1), a hematopoietic growth factor that plays an important role in the proliferation of macrophages and osteoclasts from hematopoietic stem cells. Refined mapping appeared to exclude CSF1 as the site of the mutation in the subject family.

Uses of TANGO 266 Nucleic Acids, Polypeptides, and Modulators Thereof

The TANGO 266 proteins and nucleic acid molecules of the invention have at least one "TANGO 266 activity" (also referred to herein as "TANGO 266 biological activity"). TANGO 266 activity refers to an activity exerted by a TANGO 266 protein or nucleic acid molecule on a TANGO 266 responsive cell *in vivo* or *in vitro*. Such TANGO 266 activities include at least one or more of the following activities: 1) interaction of a TANGO 266

protein with a TANGO 266-target molecule; 2) activation of a TANGO 266 target molecule; 3) modulation of cellular proliferation; 4) modulation of cellular differentiation; or 5) modulation of a signaling pathway. Thus, the TANGO 266 proteins, nucleic acids and/or modulators can be used for the treatment of a disorder characterized by aberrant TANGO 266 expression and/or an aberrant TANGO 266 activity, such as proliferative and/or differentiative disorders.

As cytokines are often found in snake venom, and due to TANGO 266's significant homology to venom protein A (VPRA), found in high abundance in the venom of the black mamba (see experimental section), TANGO 266 may be a cytokine. In the same fashion as a cytokine, TANGO 266 has been shown to play a role in the proliferation and differentiation of cells, e.g., macrophages and monocytes, and can therefore be used to treat proliferative and cell differentiation-related disorders. Such proliferative disorders include but are not limited to e.g., carcinoma, e.g., lymphoma, e.g., follicular lymphoma. Due to its ability to induce the proliferation and differentiation of white blood cell types, e.g., macrophages and monocytes, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat can be used to treat include immune disorders, e.g., viral disorders (e.g., infection by HSV), cell growth disorders, e.g., cancers (e.g., carcinoma, lymphoma, e.g., follicular lymphoma), autoimmune disorders (e.g., arthritis, graft rejection (e.g., allograft rejection), T cell disorders (e.g., AIDS)) and inflammatory disorders (e.g., bacterial infection, psoriasis, septicemia, cerebral malaria, inflammatory bowel disease, arthritis (e.g., rheumatoid arthritis, osteoarthritis), and allergic inflammatory disorders (e.g., asthma, psoriasis)).

Furthermore, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat disorders associated with leukocytes, e.g., with monocytes, macrophages, lymphocytes, and granulocytes, such as leukopenias (e.g., neutropenia, monocytopenia, lymphopenia, and granulocytopenia), leukocytosis (e.g., granulocytosis, lymphocytosis, eosinophilia, monocytosis, acute and chronic lymphadenitis), malignant lymphomas (e.g., Non-Hodgkin's lymphomas, Hodgkin's lymphomas, leukemias, agnogenic myeloid metaplasia, multiple myeloma, plasmacytoma, Waldenstrom's macroglobulinemia, heavy-chain disease, monoclonal gammopathy, histiocytoses, eosinophilic granuloma, and angioimmunoblastic lymphadenopathy).

Due to its ability to induce the proliferation and differentiation of white blood cell types, e.g., macrophages and monocytes, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat hematopoeitic disorders.

For example, hematopoeitic disorders that TANGO 266 polypeptides, nucleic acids, and/or modulators thereof can be used to treat include disorders associated with abnormal monocyte and/or macrophage function, such as impaired phagocytosis, chemotaxis, or secretion of cytokines, growth factors and acute-phase reactants, resulting from certain diseases, e.g., lysosomal storage diseases (e.g., Gaucher's disease); impaired monocyte cytokine production, for example, found in some patients with disseminated nontuberculous mycobacterial infection who are not infected with HIV; leukocyte adhesion deficiency (LAD), hyperimmunoglobulin E-recurrent infection (HIE) or Job's syndrome, Chédiak-Higashi syndrome (CHS), and chronic granulomatous diseases (CGD), certain autoimmune diseases, such as systemic lupus erythematosus and other autoimmune diseases characterized by tissue deposition of immune complexes, as seen in Sjögren's syndrome, mixed cryoglobulinemia, dermatitis herpetiformis, and chronic progressive multiple sclerosis. Also included are disorders or infections that impair mononuclear phagocyte function, for example, influenza virus infection and AIDS.

Monocyte associated disorders include monocytoses such as, for example, monocytoses associated with certain infections such as tuberculosis, brucellosis, subacute bacterial endocarditis, Rocky Mountain spotted fever, malaria, and visceral leishmaniasis (kala azar), in malignancies, leukemias, myeloproliferative syndromes, hemolytic anemias, chronic idiopathic neutropenias, and granulomatous diseases such as sarcoidosis, regional enteritis, and some collagen vascular diseases.

Other monocyte associated disorders include monocytopenias such as, for example, monocytopenias that can occur with acute infections, with stress, following administration of glucocorticoids, aplastic anemia, hairy cell leukemia, and acute myelogenous leukemia and as a direct result of administration of myelotoxic and immunosuppressive drugs.

As TANGO 266 is expressed in the spleen library, TANGO 266 nucleic acids, proteins, and/or modulators thereof can be used to modulate the proliferation, differentiation, and/or function of cells that form the spleen, e.g., cells of the splenic connective tissue, e.g., splenic smooth muscle cells and/or endothelial cells of the splenic blood vessels. TANGO 266 nucleic acids, proteins, and modulators thereof can also be used to modulate the proliferation, differentiation, and/or function of cells that are processed, e.g., regenerated or phagocytized within the spleen, e.g., erythrocytes and/or B and T lymphocytes and macrophages. Thus TANGO 266 nucleic acids, proteins, and modulators thereof can be used to treat spleen, e.g., the fetal spleen, associated diseases and disorders. Examples of splenic diseases and disorders include e.g., splenic lymphoma and/or splenomegaly, and/or

phagocytotic disorders, e.g., those inhibiting macrophage engulfinent of bacteria and viruses in the bloodstream.

As TANGO 266 is expressed in the heart, TANGO 266 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, e.g., ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease.

As TANGO 266 is expressed in the pituitary, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat disorders of the pituitary gland. The pituitary secretes such hormones as thyroid stimulating hormone (TSH), follicle stimulating hormone (FSH), adrenocotropic hormone (ACTH), and others. It controls the activity of many other endocrine glands (thyroid, ovaries, adrenal, etc.). For example, such molecules can be used to treat or modulate pituitary related disorders including, without limitation, acromegaly, Cushing's syndrome, craniopharyngiomas, Empty Sella syndrome, hypogonadism, hypopituitarism, and hypophysitis, in addition to disorders of the endocrine glands the pituitary controls.

As TANGO 266 is expressed in the thyroid, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat disorders of the thyroid gland, such as hyperthyroidism (e.g., diffuse toxic hyperplasia, toxic multinodular goiter, toxic adenoma, and acute or subacute thyroiditis), hypothyroidism (e.g., cretinism and myxedema), thyroiditis (e.g., Hashimoto's thyroiditis, subacute granulomatous thyroiditis, subacute lymphocytic thyroiditis, Riedel's thryroiditis), Graves' disease, goiter (e.g., simple diffuse goiter and multinodular goiter), or tumors (e.g., adenoma, papillary carcinoma, follicular carcinoma, medullary carcinoma, undifferentiated malignant carcinoma, Hodgkin's disease, and non-Hodgkin's lymphoma).

As TANGO 266 is expressed in adrenal tissue, e.g., in adrenal medulla and adrenal cortex, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat disorders of the adrenal cortex, such as hypoadrenalism (e.g., primary chronic or acute adrenocortical insufficiency, and secondary adrenocortical insufficiency), hyperadrenalism (Cushing's syndrome, primary hyperaldosteronism, adrenal virilism, and adrenal hyperplasia), or neoplasia (e.g., adrenal adenoma and cortical carcinoma). In another example, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat disorders of the adrenal medulla, such as neoplasms (e.g., pheochromocytomas, neuroblastomas, and ganglioneuromas).

As TANGO 266 is expressed in gonadal tissue, TANGO 266 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology,

proliferation and/or differentiation of cells in the reproductive tract, particularly in the ovaries and testis.

For example, the TANGO 266 polypeptides, nucleic acids and/or modulators thereof can be used to treat or modulate disorders associated with the testis including, without limitation, the Klinefelter syndrome (both the classic and mosaic forms), XX male syndrome, variococele, germinal cell aplasia (the Sertoli cell-only syndrome), idiopathic azoospermia or severe oligospermia, crpytochidism, and immotile cilia syndrome, or testicular cancer (primary germ cell tumors of the testis). In another example, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat testicular disorders, such as unilateral testicular enlargment (e.g., nontuberculous, granulomatous orchitis), inflammatory diseases resulting in testicular dysfunction (e.g., gonorrhea and mumps), and tumors (e.g., germ cell tumors, interstitial cell tumors, androblastoma, testicular lymphoma and adenomatoid tumors).

For example, the TANGO 266 polypeptides, nucleic acids and/or modulators thereof can be used modulate the function, morphology, proliferation and/or differentiation of the ovaries. For example, such molecules can be used to treat or modulate disorders associated with the ovaries, including, without limitation, ovarian tumors, McCune-Albright syndrome (polyostotic fibrous dysplasia). In another example, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat ovarian disorders, such as ovarian endometriosis, non-neoplastic cysts (e.g., follicular and luteal cysts and polycystic ovaries) and tumors (e.g., tumors of surface epithelium, germ cell tumors, ovarian fibroma, sex cord-stromal tumors, and ovarian cancers (e.g., metastatic carcinomas, and ovarian teratoma). For example, the TANGO 266 polypeptides, nucleic acids and/or modulators can be used in the treatment of infertility.

The TANGO 266 polypeptides, nucleic acids and/or modulators thereof can additionally be used to modulate the function, morphology, proliferation and/or differentiation of cells in the tissues of the reproductive tract other than the ovaries and testis. For example, such molecules can be used to treat or modulate disorders associated with the female reproductive tract including, without limitation, uterine disorders, e.g., hyperplasia of the endometrium, uterine cancers (e.g., uterine leiomyomoma, uterine cellular leiomyoma, leiomyosarcoma of the uterus, malignant mixed mullerian Tumor of uterus, uterine Sarcoma), and dysfunctional uterine bleeding (DUB).

As TANGO 266 is expressed in the placenta, TANGO 266 polypeptides, nucleic acids, and/or modulators thereof, can be used to treat placental disorders, such as toxemia of pregnancy (e.g., preeclampsia and eclampsia), placentitis, or spontaneous abortion.

As TANGO 266 maps to the same region as the locus for osteopetrosis, autosomal dominant, type II, and as both macrophages and osteoclasts are derived from the same progenitor cell type, e.g., monocytes, TANGO 266 polypeptides, nucleic acids and/or modulators thereof can be used to modulate the function, morphology, proliferation and/or differentiation of bone and cartilage cells, e.g., osteoclasts, osteoclasts, and chondrocytes. Thus TANGO 266 polypeptides, nucleic acids and/or modulators thereof can be used to treat bone disorders, including but not limited to bone cancer, achondroplasia, osteopetrosis (e.g., osteopetrosis, autosomal domainant, type II), myeloma, fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma, osteoporosis, and bone and/or cartilage injury due to for example, trauma (e.g., bone breakage, cartilage tearing), degeneration (e.g., osteoporosis), degeneration of joints, e.g., arthritis, e.g., osteoarthritis, and bone wearing.

TANGO 267

In another aspect, the present invention is based on the discovery of nucleic acid sequences which encode a novel family of proteins referred to herein as TANGO 267 proteins. Described herein is a human TANGO 267 (SEQ ID NO:13) nucleic acid molecule and the corresponding protein which the nucleic acid molecule encodes (SEQ ID NO:14).

TANGO 267 proteins and nucleic acid molecules comprise a family of molecules having certain conserved structural and functional features. As used herein, the term "family" is intended to mean two or more proteins or nucleic acid molecules having a common structural domain and having sufficient amino acid or nucleotide sequence identity as defined herein. Family members can be from either the same or different species. For example, a family can comprises two or more proteins of human origin, or can comprise one or more proteins of human origin and one or more of non-human origin. Members of the same family may also have common structural domains.

An TANGO 267 family member can also include a MAGE-like domain. The MAGE-like domain typically includes about 50 to 250, preferably about 75 to 225, more preferably about 120 to 200, still more preferably about 150 to 180 amino acid residues in length. The MAGE-like cytoplasmic domain typically has the following consensus sequence: [L-Xaa(6)-L-V-Xaa(2)-L-Xaa(2)-K-Xaa(n1)-E-M-L-Xaa(n2)-F-G-Xaa(2)-L-K-E-Xaa-D-Xaa(n3)-G-L-L], wherein L is leucine, Xaa is any amino acid, V is valine, K is lysine, n1 is about 2-15, preferably 5-12, and more preferably 10, E is glutamate, M is methionine, n2 is about 10-40, preferably 15-30, and more preferably 25, F is phenylalanine, G is glycine, D is aspartate, and n3 is 15-40, preferably 20-32, and more preferably 27-28.

Human TANGO 267

A sequence encoding human TANGO 267 was identified by screening a human coronary artery smooth muscle cell by EST analysis. The 2925 nucleotide human TANGO 267 sequence (Figure 17A-17C; SEQ ID NO:13) includes an open reading frame which extends from nucleotide 161 to nucleotide 2494 of SEQ ID NO:13 (SEQ ID NO:15) and encodes a 778 amino acid transmembrane protein depicted in Figure 17A-17C (SEQ ID NO:14).

In another embodiment, a human TANGO 267 clone includes comprises a 2739 nucleotide cDNA (SEQ ID NO:144). The open reading frame of this cDNA comprises nucleotides 171 to 2507 (SEQ ID NO:145), and encodes a transmembrane protein comprising the 778 amino acid sequence depicted in SEQ ID NO:14.

In one embodiment of a nucleotide sequence of human TANGO 267 the nucleotide at position 211 is a guanine (G)(SEQ ID NO:). In this embodiment, the amino acid at position 17 is glutamate (E)(SEQ ID NO:). In another embodiment of a nucleotide sequence of human TANGO 267, the nucleotide at position 211 is a cytosine (C)(SEQ ID NO:). In this embodiment, the amino acid at position 17 is aspartate (D)(SEQ ID NO:). In another embodiment of a nucleotide sequence of human TANGO 267, the nucleotide at position 223 is an adenine (A)(SEQ ID NO:). In this embodiment, the amino acid at position 21 is a glutamate (E)(SEQ ID NO:). In another embodiment of a nucleotide sequence of human TANGO 267, the nucleotide at position 223 is a cytosine (C)(SEQ ID NO:). In this embodiment, the amino acid at position 21 is aspartate (D)(SEQ ID NO:). In another embodiment of a nucleotide sequence of human TANGO 267, the nucleotide at position 256 is a guanine (G)(SEQ ID NO:). In this embodiment, the amino acid at position 32 is a glutamate (E)(SEQ ID NO:). In another embodiment of a nucleotide sequence of human TANGO 267, the nucleotide at position 256 is a cytosine (C)(SEQ ID NO:). In this embodiment, the amino acid at position 32 is aspartate (D)(SEQ ID NO:).

Another embodiment of the invention includes isolated nucleic acid molecules comprising a polynucleotide having a nucleotide sequence encoding the polypeptide having the human TANGO 267 amino acid sequence in SEQ ID NO:14, but lacking the N-terminal methionine residue. In this embodiment, the nucleotide sequence of human TANGO 267, nucleotides 164-2494 (SEQ ID NO:120), encodes the human TANGO 267 amino acid sequence from amino acids 2-778 (SEQ ID NO:121).

Human TANGO 267 protein has a molecular weight of 86.2 kD prior to post-translational modification. The presence of a methionine residue at positions 5, 27, 31, 62,

144, 205, 483, 497, 572, 589, 645, 667, and 694 of SEQ ID NO:14 indicate that there can be alternative forms of human TANGO 267 of 774 amino acids of SEQ ID NO:14, 752 amino acids of SEQ ID NO:14, 748 amino acids of SEQ ID NO:14, 717 amino acids of SEQ ID NO:14, 635 amino acids of SEQ ID NO:14, 574 amino acids of SEQ ID NO:14, 296 amino acids of SEQ ID NO:14, 282 amino acids of SEQ ID NO:14, 207 amino acids of SEQ ID NO:14, 190 amino acids of SEQ ID NO:14, 134 amino acids of SEQ ID NO:14, 112 amino acids of SEQ ID NO:14 and 83 amino acids of SEQ ID NO:14, respectively.

A clone, EpT267, which encodes human TANGO 267 was deposited with the American Type Culture Collection (ATCC™, 10801 University Boulevard, Manassas, VA 20110-2209) on March 26, 1999, and assigned Accession Number 207176. This deposit will be maintained under the terms of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure. This deposit was made merely as a convenience for those of skill in the art and is not an admission that a deposit is required under 35 U.S.C. §112.

The present invention also includes TANGO 267 proteins having a transmembrane domain. As used herein, a transmembrane domain refers to an amino acid sequence having at least about 25 to about 40 amino acid residues in length and which contains at least about 65-70% hydrophobic amino acid residues such as alanine, leucine, isoleucine, phenylalanine, proline, tyrosine, tryptophan, or valine. In a preferred embodiment, a transmembrane domain contains at least about 30-35 amino acid residues, preferably about 30-35 amino acid residues, and has at least about 60-80%, more preferably 65-75%, and more preferably at least about 68% hydrophobic residues. An example of a transmembrane domain includes from about amino acids 559 to 575 of TANGO 267 (SEQ ID NO:).

In one embodiment, human TANGO 267 includes extracellular domains at amino acids 1 to 558 of SEQ ID NO:)(SEQ ID NO:122) and amino acids 773 to 778 of SEQ ID NO:)(SEQ ID NO:123), transmembrane (TM) domains at amino acids 559 to 575 of SEQ ID NO:)(SEQ ID NO:124) and amino acids 749 to 772 of SEQ ID NO:)(SEQ ID NO:125); and a cytoplasmic domain at amino acids 576 to 748 of SEQ ID NO:)(SEQ ID NO:126).

Alternatively, in another embodiment, a human TANGO 267 protein contains an extracellular domain at amino acid residues 576 to 748 of SEQ ID NO: (SEQ ID NO:), transmembrane domains at amino acid residues 147 to 170 of SEQ ID NO: (SEQ ID NO:) and amino acid residues 749 to 772, cytoplasmic domains at amino acid residues 1 to 558 of SEQ ID NO: (SEQ ID NO:) and amino acid residues 743 to 778 of SEQ ID NO: (SEQ ID NO:).

The human gene for TANGO 267 was mapped on radiation hybrid panels to the long arm of chromosome X, in the region q12. Flanking markers for this region are WI-5587 and WI-5717. The AR (androgen receptor), MSN (moesin), and OPHN (oligophrenin 1) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome X. The gs (greasy) loci also maps to this region of the mouse chromosome. The ar (androgen receptor) and sla (sex linked anemia) genes also map to this region of the mouse chromosome.

Human TANGO 267 appears to be expressed in a wide range of tissues based on EST origin.

Human TANGO 267 protein bears similarity to a human MAGE-like protein (hepatocellular carcinoma associated gene JCL-1; GenBank Accession Numbers Z98046 and U92544) Human MAGE proteins (Kirkin et al. (1998) *APMIS* 106:665-79) are melanoma associated antigens recognized by cytotoxic T lymphocytes. It has low immunogenicity. These proteins are potentially useful targets for tumor vaccines.

Figure 18 depicts the alignment of the amino acid sequence of human TANGO 267 (SEQ ID NO:14) and human MAGE-like protein (SEQ ID NO:37). In this alignment, a (•) between the two sequences indicates an exact match.

<u>Uses of TANGO 267 Nucleic Acids, Polypeptides, and Modulators Thereof</u>

The TANGO 267 proteins and nucleic acid molecules of the invention have at least one "TANGO 267 activity" (also referred to herein as "TANGO 267 biological activity"). TANGO 267 activity refers to an activity exerted by a TANGO 267 protein or nucleic acid molecule on a TANGO 267 responsive cell *in vivo* or *in vitro*. Such TANGO 267 activities include at least one or more of the following activities: 1) interaction of a TANGO 267 protein with a TANGO 267-target molecule; 2) activation of a TANGO 267 target molecule; 3) modulation of cellular proliferation, 4) modulation of cellular differentiation; or 5) modulation of a signaling pathway. Thus, the TANGO 267 proteins, nucleic acids and/or modulators can be used for the treatment of a disorder characterized by aberrant TANGO 267 expression and/or an aberrant TANGO 267 activity, such as proliferative and/or differentiative disorders.

As TANGO 267 was originally discovered in a coronary artery smooth muscle cell by EST analysis, TANGO 267 nucleic acids, proteins, and modulators thereof can be used to treat heart disorders, e.g., ischemic heart disease, atherosclerosis, hypertension, angina pectoris, Hypertrophic Cardiomyopathy, and congenital heart disease.

In another example, because human TANGO 267 protein bears similarity to a human

MAGE-like protein (hepatocellular carcinoma associated gene JCL-1), TANGO 267 polypeptides, nucleic acids, or modulators thereof, can be used to treat hepatic (liver) disorders, such as jaundice, hepatic failure, hereditary hyperbiliruinemias (e.g., Gilbert's syndrome, Crigler-Naijar syndromes and Dubin-Johnson and Rotor's syndromes), hepatic circulatory disorders (e.g., hepatic vein thrombosis and portal vein obstruction and thrombosis) hepatitis (e.g., chronic active hepatitis, acute viral hepatitis, and toxic and drug-induced hepatitis) cirrhosis (e.g., alcoholic cirrhosis, biliary cirrhosis, and hemochromatosis), or malignant tumors (e.g., primary carcinoma, hepatoblastoma, and angiosarcoma).

Furthermore, because human TANGO 267 protein bears similarity to a human MAGE-like protein (hepatocellular carcinoma associated gene JCL-1), TANGO 216 polypeptides, nucleic acids and/or modulators thereof can also be used to modulate cell adhesion in proliferative disorders, such as cancer. Examples of types of cancers include benign tumors, neoplasms or tumors (such as carcinomas, sarcomas, adenomas or myeloid lymphoma tumors, *e.g.*, fibrosarcoma, myxosarcoma, liposarcoma, chondrosarcoma, osteogenic sarcoma, chordoma, angiosarcoma, endotheliosarcoma, lymphangiosarcoma, lymphangioendotheliosarcoma, synovioma, mesothelioma, Ewing's tumor, leimyosarcoma, rhabdotheliosarcoma, colon sarcoma, pancreatic cancer, breast cancer, ovarian cancer, prostate cancer, squamous cell carcinoma, basal cell carcinoma, adenocarcinoma, sweat gland carcinoma, sebaceous gland carcinoma, papillary carcinoma, papillary adenocarcinomas, cystadcnocarcinoma, medullary carcinoma, bronchogenic carcinoma, renal cell carcinoma, hematoma, bile duct carcinoma, melanoma, choriocarcinoma, semicoma, embryonal carcinoma, Wilms' tumor, cervical cancer, testicular tumor, lung carcinoma, small cell carcinoma, bladder carcinoma, epithelial carcinoma, glioma, astrocytoma, medulloblastoma, craniopharyngioma, ependynoma, pinealoma, hemangioblastoma, retinoblastoma), leukemias, (*e.g.* acute lymphocytic leukemia), acute myelocytic leukemia (myelolastic, promyelocytic, myelomonocytic, monocytic and erythroleukemia), chronic leukemias (chronic myelocytic (granulocytic) leukemia and chronic lymphocytic leukemia), or polycythemia vera, or lymphomas (Hodgkin's disease and non-Hodgkin's diseases), multiple myelomas and Waldenström's macroglobulinemia.

TANGO 267 could be useful as a target for tumor vaccines. Accordingly, TANGO 267 proteins (including fragments of TANGO 267) and nucleic acids and/or modulators can be used as tumor vaccines.

Tables 1, 2, 3, and 4 below provide summaries of sequence information for the TANGO molecules described herein.

TABLE 1: Summary of Human TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 Sequence Information.

Gene	cDNA	ORF	Protein	Figure	Accession No.
TANGO 216	SEQ ID NO:1	SEQ ID NO:3	SEQ ID NO:2	Figs. 1A-1C	ATCC™ 207176
TANGO 261	SEQ ID NO:4	SEQ ID NO:6	SEQ ID NO:5	Figs. 5A-5B	ATCC™ 207176
TANGO 262	SEQ ID NO:7	SEQ ID NO:9	SEQ ID NO:8	Figs 9A-9B.	ATCC™ 207176
TANGO 266	SEQ ID NO:10	SEQ ID NO:12	SEQ ID NO:11	Fig. 14.	ATCC™ 207176
TANGO 267	SEQ ID NO:13	SEQ ID NO:15	SEQ ID NO:14	Figs. 17A-17C.	ATCC™ 207176

TABLE 2: Summary of Domains of Human TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 Proteins

Protein	Signal Sequence	Mature Protein	Extracellular Domain	Transmembrane Domain	Cytoplasmic Domain
TANGO 216	aa 1-33 SEQ ID NO:25	aa 34-488 SEQ ID NO:26	aa 34 to 79 SEQ ID NO:27 aa 342 to 488 SEQ ID NO:39	aa 80-97 SEQ ID NO:28 aa 318 to 341 SEQ ID NO:40	aa 98 to 317 SEQ ID NO:29
TANGO 261	aa 1-28 SEQ ID NO:30	aa 29-252 SEQ ID NO:31			
TANGO 262	aa 1-21 SEQ ID NO:32	aa 22-226 SEQ ID NO:33			
TANGO 266	aa 1-19 SEQ ID NO:34	aa 20-105 SEQ ID NO:			
TANGO 267			aa 1 to 558 SEQ ID NO: aa 773 to 778 SEQ ID NO:	aa 559 to 575 SEQ ID NO: aa 749 to 772 SEQ ID NO:	aa 576 to 748 SEQ ID NO:

TABLE 3: Summary of Mouse TANGO 216, TANGO 261, and TANGO 262 Sequence Information.

Gene	cDNA	ORF	Protein	Figure
Mouse TANGO 216	SEQ ID NO:16	SEQ ID NO:18	SEQ ID NO:17	Figs. 2A-2B
Mouse TANGO 261	SEQ ID NO:19	SEQ ID NO:21	SEQ ID NO:20	Figs. 6A-6B
Mouse TANGO 262	SEQ ID NO:22	SEQ ID NO:24	SEQ ID NO:23	Fig. 10

TABLE 4: Summary of Domains of Mouse TANGO 216, TANGO 261, and TANGO 262 Proteins

Protein	Signal Sequence	Mature Protein	Extracellular Domain	Transmembrane Domain	Cytoplasmic Domain
Mouse TANGO 216	aa SEQ ID NO:	aa SEQ ID NO:	aa 34 to 79 SEQ ID NO:46 aa 342 to 487 SEQ ID NO:47	aa 80 to 97 SEQ ID NO:48 aa 318 to 341 SEQ ID NO:49	aa 98 to 317 SEQ ID NO:50

Various aspects of the invention are described in further detail in the following subsections.

I. Isolated Nucleic Acid Molecules

One aspect of the invention pertains to isolated nucleic acid molecules that encode a polypeptide of the invention or a biologically active portion thereof, as well as nucleic acid molecules sufficient for use as hybridization probes to identify nucleic acid molecules encoding a polypeptide of the invention and fragments of such nucleic acid molecules suitable for use as PCR primers for the amplification or mutation of nucleic acid molecules. As used herein, the term "nucleic acid molecule" is intended to include DNA molecules (e.g., cDNA or genomic DNA) and RNA molecules (e.g., mRNA) and analogs of the DNA or RNA generated using nucleotide analogs. The nucleic acid molecule can be single-stranded or double-stranded, but preferably is double-stranded DNA.

An "isolated" nucleic acid molecule is one which is separated from other nucleic acid molecules which are present in the natural source of the nucleic acid molecule. Preferably, an "isolated" nucleic acid molecule is free of sequences (preferably protein

encoding sequences) which naturally flank the nucleic acid (i.e., sequences located at the 5' and 3' ends of the nucleic acid) in the genomic DNA of the organism from which the nucleic acid is derived. For example, in various embodiments, the isolated nucleic acid molecule can contain less than about 5 kB, 4 kB, 3 kB, 2 kB, 1 kB, 0.5 kB or 0.1 kB of nucleotide sequences which naturally flank the nucleic acid molecule in genomic DNA of the cell from which the nucleic acid is derived. Moreover, an "isolated" nucleic acid molecule, such as a cDNA molecule, can be substantially free of other cellular material, or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

A nucleic acid molecule of the present invention, e.g., a nucleic acid molecule having the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or a complement thereof, can be isolated using standard molecular biology techniques and the sequence information provided herein. Using all or a portion of the nucleic acid sequences of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102 as a hybridization probe, nucleic acid molecules of the invention can be isolated using standard hybridization and cloning techniques (e.g., as described in Sambrook et al., eds., *Molecular Cloning: A Laboratory Manual, 2nd ed., Cold Spring Harbor Laboratory*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989).

A nucleic acid molecule of the invention can be amplified using cDNA, mRNA or genomic DNA as a template and appropriate oligonucleotide primers according to standard PCR amplification techniques. The nucleic acid so amplified can be cloned into an appropriate vector and characterized by DNA sequence analysis. Furthermore, oligonucleotides corresponding to all or a portion of a nucleic acid molecule of the invention can be prepared by standard synthetic techniques, e.g., using an automated DNA synthesizer.

In another preferred embodiment, an isolated nucleic acid molecule of the invention comprises a nucleic acid molecule which is a complement of the nucleotide sequence of SEQ ID NO:1. 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or a portion thereof. A nucleic acid molecule which is complementary to a given nucleotide sequence is one which is sufficiently complementary to the given nucleotide sequence that it can hybridize to the given nucleotide sequence thereby forming a stable duplex.

Moreover, a nucleic acid molecule of the invention can comprise only a portion of a nucleic acid sequence encoding a full length polypeptide of the invention for example, a fragment which can be used as a probe or primer or a fragment encoding a biologically active portion of a polypeptide of the invention. The nucleotide sequence determined from the cloning one gene allows for the generation of probes and primers designed for use in identifying and/or cloning homologues in other cell types, e.g., from other tissues, as well as homologues from other mammals. The probe/primer typically comprises substantially purified oligonucleotide. The oligonucleotide typically comprises a region of nucleotide sequence that hybridizes under stringent conditions to at least about 12, preferably about 25, more preferably about 50, 75, 100, 125, 150, 175, 200, 250, 300, 350 or 400 consecutive nucleotides of the sense or anti-sense sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or of a naturally occurring mutant of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102.

Probes based on the sequence of a nucleic acid molecule of the invention can be used to detect transcripts or genomic sequences encoding the same protein molecule encoded by a selected nucleic acid molecule. The probe comprises a label group attached thereto, e.g., a radioisotope, a fluorescent compound, an enzyme, or an enzyme co-factor. Such probes can be used as part of a diagnostic test kit for identifying cells or tissues which mis-express the protein, such as by measuring levels of a nucleic acid molecule encoding the protein in a sample of cells from a subject, e.g., detecting mRNA levels or determining whether a gene encoding the protein has been mutated or deleted.

A nucleic acid fragment encoding a biologically active portion of a polypeptide of the invention can be prepared by isolating a portion of any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, expressing the encoded portion of the polypeptide protein (e.g., by recombinant expression *in vitro*) and assessing the activity of the encoded portion of the polypeptide.

The invention further encompasses nucleic acid molecules that differ from the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102 due to degeneracy of the genetic code and thus encode the same protein as that encoded by the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21,

22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102.

In addition to the nucleotide sequences of SEQ ID Nos:3, 6, 9, 12, 15, it will be appreciated by those skilled in the art that DNA sequence polymorphisms that lead to changes in the amino acid sequence may exist within a population (e.g., the human population). Such genetic polymorphisms may exist among individuals within a population due to natural allelic variation. An allele is one of a group of genes which occur alternatively at a given genetic locus. As used herein, the phrase allelic variant refers to a nucleotide sequence which occurs at a given locus or to a polypeptide encoded by the nucleotide sequence. For example, human TANGO 216 has been mapped on radiation hybrid panels to the long arm of chromosome 4, in the region q11-13, between flanking markers GCT14E02 and jktbp-rs2, and therefore, human TANGO 216 family members can include nucleotide sequence polymorphisms (*e.g.*, nucleotide sequences that vary from SEQ ID NO:) that map to this chromosome 4 region (*i.e.*, between markers GCT14E02 and jktbp-rs2). For example, the human gene for TANGO 261 has been mapped on radiation hybrid panels to the long arm of chromosome 20, in the region q13.2-13.3, between flanking markers WI-3773 and AFMA202YB9, and therefore, human TANGO 261 family members can include nucleotide sequence polymorphisms (*e.g.*, nucleotide sequences that vary from SEQ ID NO:X) that map to this chromosome 20 region (*i.e.*, between markers WI-3773 and AFMA202YB9). For example, the human gene for TANGO 262 has been mapped on radiation hybrid panels to the long arm of chromosome 14, in the region q23-q24, between flanking markers WI-6253 and WI-5815, and therefore, human TANGO 262 family members can include nucleotide sequence polymorphisms (*e.g.*, nucleotide sequences that vary from SEQ ID NO:X) that map to this chromosome 14 region (*i.e.*, between markers WI-6253 and WI-5815). For example, the human gene for TANGO 267 was mapped on radiation hybrid panels to the long arm of chromosome X, in the region q12, between flanking markers WI-5587 and WI-5717, and therefore, human TANGO 267 family members can include nucleotide sequence polymorphisms (*e.g.*, nucleotide sequences that vary from SEQ ID NO:X) that map to this chromosome X region (*i.e.*, between markers WI-5587 and WI-5717). As used herein, the terms "gene" and "recombinant gene" refer to nucleic acid molecules comprising an open reading frame encoding a polypeptide of the invention. Such natural allelic variations can typically result in 1-5% variance in the nucleotide sequence of a given gene. Alternative alleles can be identified by sequencing the gene of interest in a number of different individuals. This can be readily carried out by

using hybridization probes to identify the same genetic locus in a variety of individuals. Any and all such nucleotide variations and resulting amino acid polymorphisms or variations that are the result of natural allelic variation and that do not alter the functional activity are intended to be within the scope of the invention.

Moreover, nucleic acid molecules encoding proteins of the invention from other species (homologues), which have a nucleotide sequence which differs from that of the human protein described herein are intended to be within the scope of the invention. Nucleic acid molecules corresponding to natural allelic variants and homologues of a cDNA of the invention can be isolated based on their identity to the human nucleic acid molecule disclosed herein using the human cDNAs, or a portion thereof, as a hybridization probe according to standard hybridization techniques under stringent hybridization conditions. For example, a cDNA encoding a soluble form of a membrane-bound protein of the invention isolated based on its hybridization to a nucleic acid molecule encoding all or part of the membrane-bound form. Likewise, a cDNA encoding a membrane-bound form can be isolated based on its hybridization to a nucleic acid molecule encoding all or part of the soluble form.

Accordingly, in another embodiment, an isolated nucleic acid molecule of the invention is at least 300 (325, 350, 375, 400, 425, 450, 500, 550, 600, 650, 700, 800, 900, 1000, or 1290) nucleotides in length and hybridizes under stringent conditions to the nucleic acid molecule comprising the nucleotide sequence, preferably the coding sequence, of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or complement thereof.

As used herein, the term "hybridizes under stringent conditions" is intended to describe conditions for hybridization and washing under which nucleotide sequences at least 60% (65%, 70%, preferably 75%) identical to each other typically remain hybridized to each other. Such stringent conditions are known to those skilled in the art and can be found in *Current Protocols in Molecular Biology*, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6. A preferred, non-limiting example of stringent hybridization conditions are hybridization in 6X sodium chloride/sodium citrate (SSC) at about 45°C, followed by one or more washes in 0.2 X SSC, 0.1% SDS at 50-65°C. Preferably, an isolated nucleic acid molecule of the invention that hybridizes under stringent conditions to the sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or complement thereof, corresponds to a naturally-occurring nucleic acid molecule. As used herein, a "naturally-occurring"

nucleic acid molecule refers to an RNA or DNA molecule having a nucleotide sequence that occurs in nature (e.g., encodes a natural protein).

In addition to naturally-occurring allelic variants of a nucleic acid molecule of the invention sequence that may exist in the population, the skilled artisan will further appreciate that changes can be introduced by mutation thereby leading to changes in the amino acid sequence of the encoded protein, without altering the biological activity of the protein. For example, one can make nucleotide substitutions leading to amino acid substitutions at "non-essential" amino acid residues. A "non-essential" amino acid residue is a residue that can be altered from the wild-type sequence without altering the biological activity, whereas an "essential" amino acid residue is required for biological activity. For example, amino acid residues that are not conserved or only semi-conserved among homologues of various species may be non-essential for activity and thus would be likely targets for alteration. Alternatively, amino acid residues that are conserved among the homologues of various species (e.g., murine and human) may be essential for activity and thus would not be likely targets for alteration.

Accordingly, another aspect of the invention pertains to nucleic acid molecules encoding a polypeptide of the invention that contain changes in amino acid residues that are not essential for activity. Such polypeptides differ in amino acid sequence from SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103, yet retain biological activity. In one embodiment, the isolated nucleic acid molecule includes a nucleotide sequence encoding a protein that includes an amino acid sequence that is at least about 45% identical, 65%, 75%, 85%, 95%, or 98% identical to the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103.

An isolated nucleic acid molecule encoding a variant protein can be created by introducing one or more nucleotide substitutions, additions or deletions into the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102 such that one or more (e.g., 1, 2, 3, 4, 5, 10, 15, 20, 30, 40, 50 or more) amino acid substitutions, additions or deletions are introduced into the encoded protein. Mutations can be introduced by standard techniques, such as site-directed mutagenesis and PCR-mediated mutagenesis. Preferably, conservative amino acid substitutions are made at one or more predicted non-essential amino acid residues. A "conservative amino acid substitution" is one in which the amino acid residue is replaced with an amino acid residue having a similar side chain.

Families of amino acid residues having similar side chains have been defined in the art. These families include amino acids with basic side chains (e.g., lysine, arginine, histidine), acidic side chains (e.g., aspartic acid, glutamic acid), uncharged polar side chains (e.g., glycine, asparagine, glutamine, serine, threonine, tyrosine, cysteine), nonpolar side chains (e.g., alanine, valine, leucine, isoleucine, proline, phenylalanine, methionine, tryptophan), beta-branched side chains (e.g., threonine, valine, isoleucine) and aromatic side chains (e.g., tyrosine, phenylalanine, tryptophan, histidine). Alternatively, mutations can be introduced randomly along all or part of the coding sequence, such as by saturation mutagenesis, and the resultant mutants can be screened for biological activity to identify mutants that retain activity. Following mutagenesis, the encoded protein can be expressed recombinantly and the activity of the protein can be determined.

In a preferred embodiment, a mutant polypeptide that is a variant of a polypeptide of the invention can be assayed for: (1) the ability to form protein:protein interactions with proteins in a signaling pathway of the polypeptide of the invention; (2) the ability to bind a ligand of the polypeptide of the invention; or (3) the ability to bind to an intracellular target protein of the polypeptide of the invention. In yet another preferred embodiment, the mutant polypeptide can be assayed for the ability to modulate cellular proliferation, cellular migration or chemotaxis, or cellular differentiation.

The present invention encompasses antisense nucleic acid molecules, i.e., molecules which are complementary to a sense nucleic acid encoding a polypeptide of the invention, e.g., complementary to the coding strand of a double-stranded cDNA molecule or complementary to an mRNA sequence. Accordingly, an antisense nucleic acid can hydrogen bond to a sense nucleic acid. The antisense nucleic acid can be complementary to an entire coding strand, or to only a portion thereof, e.g., all or part of the protein coding region (or open reading frame). An antisense nucleic acid molecule can be antisense to all or part of a non-coding region of the coding strand of a nucleotide sequence encoding a polypeptide of the invention. The non-coding regions ("5' and 3' untranslated regions") are the 5' and 3' sequences which flank the coding region and are not translated into amino acids.

An antisense oligonucleotide can be, for example, about 5, 10, 15, 20, 25, 30, 35, 40, 45 or 50 nucleotides in length. An antisense nucleic acid of the invention can be constructed using chemical synthesis and enzymatic ligation reactions using procedures known in the art. For example, an antisense nucleic acid (e.g., an antisense oligonucleotide) can be chemically synthesized using naturally occurring nucleotides or variously modified

nucleotides designed to increase the biological stability of the molecules or to increase the physical stability of the duplex formed between the antisense and sense nucleic acids, e.g., phosphorothioate derivatives and acridine substituted nucleotides can be used. Examples of modified nucleotides which can be used to generate the antisense nucleic acid include 5-fluorouracil, 5-bromouracil, 5-chlorouracil, 5-iodouracil, hypoxanthine, xanthine, 4-acetylcytosine, 5-(carboxyhydroxylmethyl) uracil, 5-carboxymethylaminomethyl-2-thiouridine, 5-carboxymethylaminomethyluracil, dihydrouracil, beta-D-galactosylqueosine, inosine, N6-isopentenyladenine, 1-methylguanine, 1-methylinosine, 2,2-dimethylguanine, 2-methyladenine, 2-methylguanine, 3-methylcytosine, 5-methylcytosine, N6-adenine, 7-methylguanine, 5-methylaminomethyluracil, 5-methoxyaminomethyl-2-thiouracil, beta-D-mannosylqueosine, 5'-methoxycarboxymethyluracil, 5-methoxyuracil, 2-methylthio-N6-isopentenyladenine, uracil-5-oxyacetic acid (v), wybutoxosine, pseudouracil, queosine, 2-thiocytosine, 5-methyl-2-thiouracil, 2-thiouracil, 4-thiouracil, 5-methyluracil, uracil-5-oxyacetic acid methylester, uracil-5-oxyacetic acid (v), 5-methyl-2-thiouracil, 3-(3-amino-3-N-2-carboxypropyl) uracil, (acp3)w, and 2,6-diaminopurine. Alternatively, the antisense nucleic acid can be produced biologically using an expression vector into which a nucleic acid has been subcloned in an antisense orientation (i.e., RNA transcribed from the inserted nucleic acid will be of an antisense orientation to a target nucleic acid of interest, described further in the following subsection).

The antisense nucleic acid molecules of the invention are typically administered to a subject or generated *in situ* such that they hybridize with or bind to cellular mRNA and/or genomic DNA encoding a selected polypeptide of the invention to thereby inhibit expression, e.g., by inhibiting transcription and/or translation. The hybridization can be by conventional nucleotide complementarity to form a stable duplex, or, for example, in the case of an antisense nucleic acid molecule which binds to DNA duplexes, through specific interactions in the major groove of the double helix. An example of a route of administration of antisense nucleic acid molecules of the invention includes direct injection at a tissue site. Alternatively, antisense nucleic acid molecules can be modified to target selected cells and then administered systemically. For example, for systemic administration, antisense molecules can be modified such that they specifically bind to receptors or antigens expressed on a selected cell surface, e.g., by linking the antisense nucleic acid molecules to peptides or antibodies which bind to cell surface receptors or antigens. The antisense nucleic acid molecules can also be delivered to cells using the vectors described herein. To achieve sufficient intracellular concentrations of the antisense molecules, vector constructs

in which the antisense nucleic acid molecule is placed under the control of a strong pol II or pol III promoter are preferred.

An antisense nucleic acid molecule of the invention can be an α-anomeric nucleic acid molecule. An α-anomeric nucleic acid molecule forms specific double-stranded hybrids with complementary RNA in which, contrary to the usual β-units, the strands run parallel to each other (Gaultier et al. (1987) *Nucleic Acids Res.* 15:6625-6641). The antisense nucleic acid molecule can also comprise a 2'-o-methylribonucleotide (Inoue et al. (1987) *Nucleic Acids Res.* 15:6131-6148) or a chimeric RNA-DNA analogue (Inoue et al. (1987) *FEBS Lett.* 215:327-330).

The invention also encompasses ribozymes. Ribozymes are catalytic RNA molecules with ribonuclease activity which are capable of cleaving a single-stranded nucleic acid, such as an mRNA, to which they have a complementary region. Thus, ribozymes (e.g., hammerhead ribozymes (described in Haselhoff and Gerlach (1988) *Nature* 334:585-591)) can be used to catalytically cleave mRNA transcripts to thereby inhibit translation of the protein encoded by the mRNA. A ribozyme having specificity for a nucleic acid molecule encoding a polypeptide of the invention can be designed based upon the nucleotide sequence of a cDNA disclosed herein. For example, a derivative of a *Tetrahymena* L-19 IVS RNA can be constructed in which the nucleotide sequence of the active site is complementary to the nucleotide sequence to be cleaved in a Cech et al. U.S. Patent No. 4,987,071; and Cech et al. U.S. Patent No. 5,116,742. Alternatively, an mRNA encoding a polypeptide of the invention can be used to select a catalytic RNA having a specific ribonuclease activity from a pool of RNA molecules. *See, e.g.*, Bartel and Szostak (1993) *Science* 261:1411-1418.

The invention also encompasses nucleic acid molecules which form triple helical structures. For example, expression of a polypeptide of the invention can be inhibited by targeting nucleotide sequences complementary to the regulatory region of the gene encoding the polypeptide (e.g., the promoter and/or enhancer) to form triple helical structures that prevent transcription of the gene in target cells. *See generally* Helene (1991) *Anticancer Drug Des.* 6(6):569-84; Helene (1992) *Ann. N.Y. Acad. Sci.* 660:27-36; and Maher (1992) *Bioassays* 14(12):807-15.

In various embodiments, the nucleic acid molecules of the invention can be modified at the base moiety, sugar moiety or phosphate backbone to improve, e.g., the stability, hybridization, or solubility of the molecule. For example, the deoxyribose phosphate backbone of the nucleic acids can be modified to generate peptide nucleic acids (*see* Hyrup

et al. (1996) *Bioorganic & Medicinal Chemistry* 4(1): 5-23). As used herein, the terms "peptide nucleic acids" or "PNAs" refer to nucleic acid mimics, e.g., DNA mimics, in which the deoxyribose phosphate backbone is replaced by a pseudopeptide backbone and only the four natural nucleobases are retained. The neutral backbone of PNAs has been shown to allow for specific hybridization to DNA and RNA under conditions of low ionic strength. The synthesis of PNA oligomers can be performed using standard solid phase peptide synthesis protocols as described in Hyrup et al. (1996), *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 14670-675.

PNAs can be used in therapeutic and diagnostic applications. For example, PNAs can be used as antisense or antigene agents for sequence-specific modulation of gene expression by, e.g., inducing transcription or translation arrest or inhibiting replication. PNAs can also be used, e.g., in the analysis of single base pair mutations in a gene by, e.g., PNA directed PCR clamping; as artificial restriction enzymes when used in combination with other enzymes, e.g., S1 nucleases (Hyrup (1996), *supra*; or as probes or primers for DNA sequence and hybridization (Hyrup (1996), *supra*; Perry-O'Keefe et al. (1996) *Proc. Natl. Acad. Sci. USA* 93: 14670-675).

In another embodiment, PNAs can be modified, e.g., to enhance their stability or cellular uptake, by attaching lipophilic or other helper groups to PNA, by the formation of PNA-DNA chimeras, or by the use of liposomes or other techniques of drug delivery known in the art. For example, PNA-DNA chimeras can be generated which may combine the advantageous properties of PNA and DNA. Such chimeras allow DNA recognition enzymes, e.g., RNAse H and DNA polymerases, to interact with the DNA portion while the PNA portion would provide high binding affinity and specificity. PNA-DNA chimeras can be linked using linkers of appropriate lengths selected in terms of base stacking, number of bonds between the nucleobases, and orientation (Hyrup (1996), *supra*). The synthesis of PNA-DNA chimeras can be performed as described in Hyrup (1996), *supra,* and Finn et al. (1996) *Nucleic Acids Res.* 24(17):3357-63. For example, a DNA chain can be synthesized on a solid support using standard phosphoramidite coupling chemistry and modified nucleoside analogs. Compounds such as 5'-(4-methoxytrityl)amino-5'-deoxy-thymidine phosphoramidite can be used as a link between the PNA and the 5' end of DNA (Mag et al. (1989) *Nucleic Acids Res.* 17:5973-88). PNA monomers are then coupled in a stepwise manner to produce a chimeric molecule with a 5' PNA segment and a 3' DNA segment (Finn et al. (1996) *Nucleic Acids Res.* 24(17):3357-63). Alternatively, chimeric molecules can be synthesized with a 5' DNA segment and a 3' PNA segment (Peterser et al. (1975) *Bioorganic Med. Chem. Lett.* 5:1119-11124).

In other embodiments, the oligonucleotide may include other appended groups such as peptides (e.g., for targeting host cell receptors *in vivo*), or agents facilitating transport across the cell membrane (*see, e.g.*, Letsinger et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6553-6556; Lemaitre et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:648-652; PCT Publication No. W0 88/09810) or the blood-brain barrier (*see, e.g.*, PCT Publication No. W0 89/10134). In addition, oligonucleotides can be modified with hybridization-triggered cleavage agents (*see, e.g.*, Krol et al. (1988) *Bio/Techniques* 6:958-976) or intercalating agents (*see, e.g.*, Zon (1988) *Pharm. Res.* 5:539-549). To this end, the oligonucleotide may be conjugated to another molecule, e.g., a peptide, hybridization triggered cross-linking agent, transport agent, hybridization-triggered cleavage agent, etc.

II. Isolated Proteins and Antibodies

One aspect of the invention pertains to isolated proteins, and biologically active portions thereof, as well as polypeptide fragments suitable for use as immunogens to raise antibodies directed against a polypeptide of the invention. In one embodiment, the native polypeptide can be isolated from cells or tissue sources by an appropriate purification scheme using standard protein purification techniques. In another embodiment, polypeptides of the invention are produced by recombinant DNA techniques. Alternative to recombinant expression, a polypeptide of the invention can be synthesized chemically using standard peptide synthesis techniques.

An "isolated" or "purified" protein or biologically active portion thereof is substantially free of cellular material or other contaminating proteins from the cell or tissue source from which the protein is derived, or substantially free of chemical precursors or other chemicals when chemically synthesized. The language "substantially free of cellular material" includes preparations of protein in which the protein is separated from cellular components of the cells from which it is isolated or recombinantly produced. Thus, protein that is substantially free of cellular material includes preparations of protein having less than about 30%, 20%, 10%, or 5% (by dry weight) of heterologous protein (also referred to herein as a "contaminating protein"). When the protein or biologically active portion thereof is recombinantly produced, it is also preferably substantially free of culture medium, i.e., culture medium represents less than about 20%, 10%, or 5% of the volume of the protein preparation. When the protein is produced by chemical synthesis, it is preferably substantially free of chemical precursors or other chemicals, i.e., it is separated from chemical precursors or other chemicals which are involved in the synthesis of the protein.

Accordingly such preparations of the protein have less than about 30%, 20%, 10%, 5% (by dry weight) of chemical precursors or compounds other than the polypeptide of interest.

Biologically active portions of a polypeptide of the invention include polypeptides comprising amino acid sequences sufficiently identical to or derived from the amino acid sequence of the protein (e.g., the amino acid sequence shown in any of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103), which include fewer amino acids than the full length protein, and exhibit at least one activity of the corresponding full-length protein. Typically, biologically active portions comprise a domain or motif with at least one activity of the corresponding protein. A biologically active portion of a protein of the invention can be a polypeptide which is, for example, 10, 25, 50, 100 or more amino acids in length. Moreover, other biologically active portions, in which other regions of the protein are deleted, can be prepared by recombinant techniques and evaluated for one or more of the functional activities of the native form of a polypeptide of the invention.

Preferred polypeptides have the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103. Other useful proteins are substantially identical (e.g., at least about 45%, preferably 55%, 65%, 75%, 85%, 95%, or 99%) to any of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103 and retain the functional activity of the protein of the corresponding naturally-occurring protein yet differ in amino acid sequence due to natural allelic variation or mutagenesis.

To determine the percent identity of two amino acid sequences or of two nucleic acids, the sequences are aligned for optimal comparison purposes (e.g., gaps can be introduced in the sequence of a first amino acid or nucleic acid sequence for optimal alignment with a second amino or nucleic acid sequence). The amino acid residues or nucleotides at corresponding amino acid positions or nucleotide positions are then compared. When a position in the first sequence is occupied by the same amino acid residue or nucleotide as the corresponding position in the second sequence, then the molecules are identical at that position. The percent identity between the two sequences is a function of the number of identical positions shared by the sequences (i.e., % identity = # of identical positions/total # of positions (e.g., overlapping positions) x 100). In one embodiment the two sequences are the same length.

The determination of percent identity between two sequences can be accomplished using a mathematical algorithm. A preferred, non-limiting example of a mathematical algorithm utilized for the comparison of two sequences is the algorithm of Karlin and Altschul (1990) *Proc. Natl. Acad. Sci. USA* 87:2264-2268, modified as in Karlin and Altschul (1993) *Proc. Natl. Acad. Sci. USA* 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul, et al. (1990) *J. Mol. Biol.* 215:403-410. BLAST nucleotide searches can be performed with the NBLAST program, score = 100, wordlength = 12 to obtain nucleotide sequences homologous to a nucleic acid molecules of the invention. BLAST protein searches can be performed with the XBLAST program, score = 50, wordlength = 3 to obtain amino acid sequences homologous to a protein molecules of the invention. To obtain gapped alignments for comparison purposes, Gapped BLAST can be utilized as described in Altschul et al. (1997) *Nucleic Acids Res.* 25:3389-3402. Alternatively, PSI-Blast can be used to perform an iterated search which detects distant relationships between molccules. *Id.* When utilizing BLAST, Gapped BLAST, and PSI-Blast programs, the default parameters of the respective programs (e.g., XBLAST and NBLAST) can be used. *See* http://www.ncbi.nlm.nih.gov. Another preferred, non-limiting example of a mathematical algorithm utilized for the comparison of sequences is the algorithm of Myers and Miller, (1988) *CABIOS* 4:11-17. Such an algorithm is incorporated into the ALIGN program (version 2.0) which is part of the GCG sequence alignment software package. When utilizing the ALIGN program for comparing amino acid sequences, a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4 can be used.

The percent identity between two sequences can be determined using techniques similar to those described above, with or without allowing gaps. In calculating percent identity, only exact matches are counted.

The invention also provides chimeric or fusion proteins. As used herein, a "chimeric protein" or "fusion protein" comprises all or part (preferably biologically active) of a polypeptide of the invention operably linked to a heterologous polypeptide (i.e., a polypeptide other than the same polypeptide of the invention). Within the fusion protein, the term "operably linked" is intended to indicate that the polypeptide of the invention and the heterologous polypeptide are fused in-frame to each other. The heterologous polypeptide can be fused to the N-terminus or C-terminus of the polypeptide of the invention.

One useful fusion protein is a GST fusion protein in which the polypeptide of the invention is fused to the C-terminus of GST sequences. Such fusion proteins can facilitate the purification of a recombinant polypeptide of the invention.

In another embodiment, the fusion protein contains a heterologous signal sequence at its N-terminus. For example, the native signal sequence of a polypeptide of the invention can be removed and replaced with a signal sequence from another protein. For example, the gp67 secretory sequence of the baculovirus envelope protein can be used as a heterologous signal sequence (*Current Protocols in Molecular Biology,* Ausubel et al., eds., John Wiley & Sons, 1992). Other examples of eukaryotic heterologous signal sequences include the secretory sequences of melittin and human placental alkaline phosphatase (Stratagene; La Jolla, California). In yet another example, useful prokaryotic heterologous signal sequences include the phoA secretory signal (Sambrook et al., *supra*) and the protein A secretory signal (Pharmacia Biotech; Piscataway, New Jersey).

In yet another embodiment, the fusion protein is an immunoglobulin fusion protein in which all or part of a polypeptide of the invention is fused to sequences derived from a member of the immunoglobulin protein family. The immunoglobulin fusion proteins of the invention can be incorporated into pharmaceutical compositions and administered to a subject to inhibit an interaction between a ligand (soluble or membrane-bound) and a protein on the surface of a cell (receptor), to thereby suppress signal transduction *in vivo*. The immunoglobulin fusion protein can be used to affect the bioavailability of a cognate ligand of a polypeptide of the invention. Inhibition of ligand/receptor interaction may be useful therapeutically, both for treating proliferative and differentiative disorders and for modulating (e.g. promoting or inhibiting) cell survival. Moreover, the immunoglobulin fusion proteins of the invention can be used as immunogens to produce antibodies directed against a polypeptide of the invention in a subject, to purify ligands and in screening assays to identify molecules which inhibit the interaction of receptors with ligands.

Chimeric and fusion proteins of the invention can be produced by standard recombinant DNA techniques. In another embodiment, the fusion gene can be synthesized by conventional techniques including automated DNA synthesizers. Alternatively, PCR amplification of gene fragments can be carried out using anchor primers which give rise to complementary overhangs between two consecutive gene fragments which can subsequently be annealed and reamplified to generate a chimeric gene sequence (*see, e.g.*, Ausubel et al., *supra*). Moreover, many expression vectors are commercially available that already encode a fusion moiety (e.g., a GST polypeptide). A nucleic acid encoding a polypeptide of the

invention can be cloned into such an expression vector such that the fusion moiety is linked in-frame to the polypeptide of the invention.

A signal sequence of a polypeptide of the invention (SEQ ID NO:25, 30, 32 or 34) can be used to facilitate secretion and isolation of the secreted protein or other proteins of interest. Signal sequences are typically characterized by a core of hydrophobic amino acids which are generally cleaved from the mature protein during secretion in one or more cleavage events. Such signal peptides contain processing sites that allow cleavage of the signal sequence from the mature proteins as they pass through the secretory pathway. Thus, the invention pertains to the described polypeptides having a signal sequence, as well as to the signal sequence itself and to the polypeptide in the absence of the signal sequence (i.e., the cleavage products). In one embodiment, a nucleic acid sequence encoding a signal sequence of the invention can be operably linked in an expression vector to a protein of interest, such as a protein which is ordinarily not secreted or is otherwise difficult to isolate. The signal sequence directs secretion of the protein, such as from a eukaryotic host into which the expression vector is transformed, and the signal sequence is subsequently or concurrently cleaved. The protein can then be readily purified from the extracellular medium by art recognized methods. Alternatively, the signal sequence can be linked to the protein of interest using a sequence which facilitates purification, such as with a GST domain.

In another embodiment, the signal sequences of the present invention can be used to identify regulatory sequences, e.g., promoters, enhancers, repressors. Since signal sequences are the most amino-terminal sequences of a peptide, it is expected that the nucleic acids which flank the signal sequence on its amino-terminal side will be regulatory sequences which affect transcription. Thus, a nucleotide sequence which encodes all or a portion of a signal sequence can be used as a probe to identify and isolate signal sequences and their flanking regions, and these flanking regions can be studied to identify regulatory elements therein.

The present invention also pertains to variants of the polypeptides of the invention. Such variants have an altered amino acid sequence which can function as either agonists (mimetics) or as antagonists. Variants can be generated by mutagenesis, e.g., discrete point mutation or truncation. An agonist can retain substantially the same, or a subset, of the biological activities of the naturally occurring form of the protein. An antagonist of a protein can inhibit one or more of the activities of the naturally occurring form of the protein by, for example, competitively binding to a downstream or upstream member of a

cellular signaling cascade which includes the protein of interest. Thus, specific biological effects can be elicited by treatment with a variant of limited function. Treatment of a subject with a variant having a subset of the biological activities of the naturally occurring form of the protein can have fewer side effects in a subject relative to treatment with the naturally occurring form of the protein.

Variants of a protein of the invention which function as either agonists (mimetics) or as antagonists can be identified by screening combinatorial libraries of mutants, e.g., truncation mutants, of the protein of the invention for agonist or antagonist activity. In one embodiment, a variegated library of variants is generated by combinatorial mutagenesis at the nucleic acid level and is encoded by a variegated gene library. A variegated library of variants can be produced by, for example, enzymatically ligating a mixture of synthetic oligonucleotides into gene sequences such that a degenerate set of potential protein sequences is expressible as individual polypeptides, or alternatively, as a set of larger fusion proteins (e.g., for phage display). There are a variety of methods which can be used to produce libraries of potential variants of the polypeptides of the invention from a degenerate oligonucleotide sequence. Methods for synthesizing degenerate oligonucleotides are known in the art (*see, e.g.,* Narang (1983) *Tetrahedron* 39:3; Itakura et al. (1984) *Annu. Rev. Biochem.* 53:323; Itakura et al. (1984) *Science* 198:1056; Ike et al. (1983) *Nucleic Acid Res.* 11:477).

In addition, libraries of fragments of the coding sequence of a polypeptide of the invention can be used to generate a variegated population of polypeptides for screening and subsequent selection of variants. For example, a library of coding sequence fragments can be generated by treating a double stranded PCR fragment of the coding sequence of interest with a nuclease under conditions wherein nicking occurs only about once per molecule, denaturing the double stranded DNA, renaturing the DNA to form double stranded DNA which can include sense/antisense pairs from different nicked products, removing single stranded portions from reformed duplexes by treatment with S1 nuclease, and ligating the resulting fragment library into an expression vector. By this method, an expression library can be derived which encodes N-terminal and internal fragments of various sizes of the protein of interest.

Several techniques are known in the art for screening gene products of combinatorial libraries made by point mutations or truncation, and for screening cDNA libraries for gene products having a selected property. The most widely used techniques, which are amenable to high through-put analysis, for screening large gene libraries typically include cloning the

gene library into replicable expression vectors, transforming appropriate cells with the resulting library of vectors, and expressing the combinatorial genes under conditions in which detection of a desired activity facilitates isolation of the vector encoding the gene whose product was detected. Recursive ensemble mutagenesis (REM), a technique which enhances the frequency of functional mutants in the libraries, can be used in combination with the screening assays to identify variants of a protein of the invention (Arkin and Yourvan (1992) *Proc. Natl. Acad. Sci. USA 89*:7811-7815; Delgrave et al. (1993) *Protein Engineering* 6(3):327-331).

An isolated polypeptide of the invention, or a fragment thereof, can be used as an immunogen to generate antibodies using standard techniques for polyclonal and monoclonal antibody preparation. The full-length polypeptide or protein can be used or, alternatively, the invention provides antigenic peptide fragments for use as immunogens. The antigenic peptide of a protein of the invention comprises at least 8 (preferably 10, 15, 20, or 30) amino acid residues of the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103, and encompasses an epitope of the protein such that an antibody raised against the peptide forms a specific immune complex with the protein.

Preferred epitopes encompassed by the antigenic peptide are regions that are located on the surface of the protein, e.g., hydrophilic regions. Figures 3, 7, 11, and 15 represent hydrophobicity plots of the proteins of the invention. These plots or similar analyses can be used to identify hydrophilic regions.

An immunogen typically is used to prepare antibodies by immunizing a suitable subject, (e.g., rabbit, goat, mouse or other mammal). An appropriate immunogenic preparation can contain, for example, recombinantly expressed or chemically synthesized polypeptide. The preparation can further include an adjuvant, such as Freund's complete or incomplete adjuvant, or similar immunostimulatory agent.

Accordingly, another aspect of the invention pertains to antibodies directed against a polypeptide of the invention. The term "antibody" as used herein refers to immunoglobulin molecules and immunologically active portions of immunoglobulin molecules, i.e., molecules that contain an antigen binding site which specifically binds an antigen, such as a polypeptide of the invention, *e.g.*, an epitope of a polypeptide of the invention. A molecule which specifically binds to a given polypeptide of the invention is a molecule which binds the polypeptide, but does not substantially bind other molecules in a sample, e.g., a biological sample, which naturally contains the polypeptide. Examples of

immunologically active portions of immunoglobulin molecules include F(ab) and $F(ab')_2$ fragments which can be generated by treating the antibody with an enzyme such as pepsin. The invention provides polyclonal and monoclonal antibodies. The term "monoclonal antibody" or "monoclonal antibody composition", as used herein, refers to a population of antibody molecules that contain only one species of an antigen binding site capable of immunoreacting with a particular epitope.

Polyclonal antibodies can be prepared as described above by immunizing a suitable subject with a polypeptide of the invention as an immunogen. Preferred polyclonal antibody compositions are ones that have been selected for antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred polyclonal antibody preparations are ones that contain only antibodies directed against a polypeptide or polypeptides of the invention. Particularly preferred immunogen compositions are those that contain no other human proteins such as, for example, immunogen compositions made using a non-human host cell for recombinant expression of a polypeptide of the invention. In such a manner, the only human epitope or epitopes recognized by the resulting antibody compositions raised against this immunogen will be present as part of a polypeptide or polypeptides of the invention.

The antibody titer in the immunized subject can be monitored over time by standard techniques, such as with an enzyme linked immunosorbent assay (ELISA) using immobilized polypeptide. If desired, the antibody molecules can be isolated from the mammal (e.g., from the blood) and further purified by well-known techniques, such as protein A chromatography to obtain the IgG fraction. Alternatively, antibodies specific for a protein or polypeptide of the invention can be selected for (*e.g.*, partially purified) or purified by, *e.g.*, affinity chromatography. For example, a recombinantly expressed and purified (or partially purified) protein of the invention is produced as described herein, and covalently or non-covalently coupled to a solid support such as, for example, a chromatography column. The column can then be used to affinity purify antibodies specific for the proteins of the invention from a sample containing antibodies directed against a large number of different epitopes, thereby generating a substantially purified antibody composition, *i.e.*, one that is substantially free of contaminating antibodies. By a substantially purified antibody composition is meant, in this context, that the antibody sample contains at most only 30% (by dry weight) of contaminating antibodies directed against epitopes other than those on the desired protein or polypeptide of the invention, and preferably at most 20%, yet more preferably at most 10%, and most preferably at most 5%

(by dry weight) of the sample is contaminating antibodies. A purified antibody composition means that at least 99% of the antibodies in the composition are directed against the desired protein or polypeptide of the invention.

At an appropriate time after immunization, e.g., when the specific antibody titers are highest, antibody-producing cells can be obtained from the subject and used to prepare monoclonal antibodies by standard techniques, such as the hybridoma technique originally described by Kohler and Milstein (1975) *Nature* 256:495-497, the human B cell hybridoma technique (Kozbor et al. (1983) *Immunol. Today* 4:72), the EBV-hybridoma technique (Cole et al. (1985), *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, Inc., pp. 77-96) or trioma techniques. The technology for producing hybridomas is well known (*see generally Current Protocols in Immunology* (1994) Coligan et al. (eds.) John Wiley & Sons, Inc., New York, NY). Hybridoma cells producing a monoclonal antibody of the invention are detected by screening the hybridoma culture supernatants for antibodies that bind the polypeptide of interest, e.g., using a standard ELISA assay.

Alternative to preparing monoclonal antibody-secreting hybridomas, a monoclonal antibody directed against a polypeptide of the invention can be identified and isolated by screening a recombinant combinatorial immunoglobulin library (e.g., an antibody phage display library) with the polypeptide of interest. Kits for generating and screening phage display libraries are commercially available (e.g., the Pharmacia *Recombinant Phage Antibody System*, Catalog No. 27-9400-01; and the Stratagene *SurfZAP Phage Display Kit*, Catalog No. 240612). Additionally, examples of methods and reagents particularly amenable for use in generating and screening antibody display library can be found in, for example, U.S. Patent No. 5,223,409; PCT Publication No. WO 92/18619; PCT Publication No. WO 91/17271; PCT Publication No. WO 92/20791; PCT Publication No. WO 92/15679; PCT Publication No. WO 93/01288; PCT Publication No. WO 92/01047; PCT Publication No. WO 92/09690; PCT Publication No. WO 90/02809; Fuchs et al. (1991) *Bio/Technology* 9:1370-1372; Hay et al. (1992) *Hum. Antibod. Hybridomas* 3:81-85; Huse et al. (1989) *Science* 246:1275-1281; Griffiths et al. (1993) *EMBO J.* 12:725-734.

Additionally, recombinant antibodies, such as chimeric and humanized monoclonal antibodies, comprising both human and non-human portions, which can be made using standard recombinant DNA techniques, are within the scope of the invention. A chimeric antibody is a molecule in which different portions are derived from different animal species, such as those having a variable region derived from a murine mAb and a human immunoglobulin constant region. (See, *e.g.*, Cabilly et al., U.S. Patent No. 4,816,567; and

Boss et al., U.S. Patent No. 4,816397, which are incorporated herein by reference in their entirety.) Humanized antibodies are antibody molecules from non-human species having one or more complementarily determining regions (CDRs) from the non-human species and a framework region from a human immunoglobulin molecule. (See, *e.g.*, Queen, U.S. Patent No. 5,585,089, which is incorporated herein by reference in its entirety.) Such chimeric and humanized monoclonal antibodies can be produced by recombinant DNA techniques known in the art, for example using methods described in PCT Publication No. WO 87/02671; European Patent Application 184,187; European Patent Application 171,496; European Patent Application 173,494; PCT Publication No. WO 86/01533; U.S. Patent No 4,816,567; European Patent Application 125,023; Better et al. (1988) *Science* 240:1041-1043; Liu et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:3439-3443; Liu et al. (1987) *J. Immunol.* 139:3521-3526; Sun et al. (1987) *Proc. Natl. Acad. Sci. USA* 84:214-218; Nishimura et al. (1987) *Canc. Res.* 47:999-1005; Wood et al. (1985) *Nature* 314:446-449; and Shaw et al. (1988) *J. Natl. Cancer Inst.* 80:1553-1559); Morrison (1985) *Science* 229:1202-1207; Oi et al. (1986) *Bio/Techniques* 4:214; U.S. Patent 5,225,539; Jones et al. (1986) *Nature* 321:552-525; Verhoeyan et al. (1988) *Science* 239:1534; and Beidler et al. (1988) *J. Immunol.* 141:4053-4060.

Completely human antibodies are particularly desirable for therapeutic treatment of human patients. Such antibodies can be produced, for example, using transgenic mice which are incapable of expressing endogenous immunoglobulin heavy and light chains genes, but which can express human heavy and light chain genes. The transgenic mice are immunized in the normal fashion with a selected antigen, e.g., all or a portion of a polypeptide of the invention. Monoclonal antibodies directed against the antigen can be obtained using conventional hybridoma technology. The human immunoglobulin transgenes harbored by the transgenic mice rearrange during B cell differentiation, and subsequently undergo class switching and somatic mutation. Thus, using such a technique, it is possible to produce therapeutically useful IgG, IgA and IgE antibodies. For an overview of this technology for producing human antibodies, see Lonberg and Huszar (1995, *Int. Rev. Immunol.* 13:65-93). For a detailed discussion of this technology for producing human antibodies and human monoclonal antibodies and protocols for producing such antibodies, *see, e.g.*, U.S. Patent 5,625,126; U.S. Patent 5,633,425; U.S. Patent 5,569,825; U.S. Patent 5,661,016; and U.S. Patent 5,545,806. In addition, companies such as Abgenix, Inc. (Freemont, CA), can be engaged to provide human antibodies directed against a selected antigen using technology similar to that described above.

Completely human antibodies which recognize a selected epitope can be generated using a technique referred to as "guided selection." In this approach a selected non-human monoclonal antibody, e.g., a murine antibody, is used to guide the selection of a completely human antibody recognizing the same epitope. (Jespers et al. (1994) *Bio/technology* 12:899-903).

An antibody directed against a polypeptide of the invention (e.g., monoclonal antibody) can be used to isolate the polypeptide by standard techniques, such as affinity chromatography or immunoprecipitation. Moreover, such an antibody can be used to detect the protein (e.g., in a cellular lysate or cell supernatant) in order to evaluate the abundance and pattern of expression of the polypeptide. The antibodies can also be used diagnostically to monitor protein levels in tissue as part of a clinical testing procedure, e.g., to, for example, determine the efficacy of a given treatment regimen. Detection can be facilitated by coupling the antibody to a detectable substance. Examples of detectable substances include various enzymes, prosthetic groups, fluorescent materials, luminescent materials, bioluminescent materials, and radioactive materials. Examples of suitable enzymes include horseradish peroxidase, alkaline phosphatase, β-galactosidase, or acetylcholinesterase; examples of suitable prosthetic group complexes include streptavidin/biotin and avidin/biotin; examples of suitable fluorescent materials include umbelliferone, fluorescein, fluorescein isothiocyanate, rhodamine, dichlorotriazinylamine fluorescein, dansyl chloride or phycoerythrin; an example of a luminescent material includes luminol; examples of bioluminescent materials include luciferase, luciferin, and aequorin, and examples of suitable radioactive material include ^{125}I, ^{131}I, ^{35}S or ^{3}H.

Further, an antibody (or fragment thereof) can be conjugated to a therapeutic moiety such as a cytotoxin, a therapeutic agent or a radioactive metal ion. A cytotoxin or cytotoxic agent includes any agent that is detrimental to cells. Examples include taxol, cytochalasin B, gramicidin D, ethidium bromide, emetine, mitomycin, etoposide, tenoposide, vincristine, vinblastine, colchicin, doxorubicin, daunorubicin, dihydroxy anthracin dione, mitoxantrone, mithramycin, actinomycin D, 1-dehydrotestosterone, glucocorticoids, procaine, tetracaine, lidocaine, propranolol, and puromycin and analogs or homologs thereof. Therapeutic agents include, but are not limited to, antimetabolites (*e.g.*, methotrexate, 6-mercaptopurine, 6-thioguanine, cytarabine, 5-fluorouracil decarbazine), alkylating agents (*e.g.*, mechlorethamine, thioepa chlorambucil, melphalan, carmustine (BSNU) and lomustine (CCNU), cyclothosphamide, busulfan, dibromomannitol, streptozotocin, mitomycin C, and cis-dichlorodiamine platinum (II) (DDP) cisplatin), anthracyclines (*e.g.*, daunorubicin (formerly daunomycin) and doxorubicin), antibiotics (*e.g.*, dactinomycin (formerly

actinomycin), bleomycin, mithramycin, and anthramycin (AMC)), and anti-mitotic agents (*e.g.*, vincristine and vinblastine).

The conjugates of the invention can be used for modifying a given biological response, the drug moiety is not to be construed as limited to classical chemical therapeutic agents. For example, the drug moiety may be a protein or polypeptide possessing a desired biological activity. Such proteins may include, for example, a toxin such as abrin, ricin A, pseudomonas exotoxin, or diphtheria toxin; a protein such as tumor necrosis factor, .alpha.-interferon, .beta.-interferon, nerve growth factor, platelet derived growth factor, tissue plasminogen activator; or, biological response modifiers such as, for example, lymphokines, interleukin-1 ("IL-1"), interleukin-2 ("IL-2"), interleukin-6 ("IL-6"), granulocyte macrophase colony stimulating factor ("GM-CSF"), granulocyte colony stimulating factor ("G-CSF"), or other growth factors.

Techniques for conjugating such therapeutic moiety to antibodies are well known, see, *e.g.*, Arnon et al., "Monoclonal Antibodies For Immunotargeting Of Drugs In Cancer Therapy", in Monoclonal Antibodies And Cancer Therapy, Reisfeld et al. (eds.), pp. 243-56 (Alan R. Liss, Inc. 1985); Hellstrom et al., "Antibodies For Drug Delivery", in Controlled Drug Delivery (2nd Ed.), Robinson et al. (eds.), pp. 623-53 (Marcel Dekker, Inc. 1987); Thorpe, "Antibody Carriers Of Cytotoxic Agents In Cancer Therapy: A Review", in Monoclonal Antibodies '84: Biological And Clinical Applications, Pinchera et al. (eds.), pp. 475-506 (1985); "Analysis, Results, And Future Prospective Of The Therapeutic Use Of Radiolabeled Antibody In Cancer Therapy", in Monoclonal Antibodies For Cancer Detection And Therapy, Baldwin et al. (eds.), pp. 303-16 (Academic Press 1985), and Thorpe et al., "The Preparation And Cytotoxic Properties Of Antibody-Toxin Conjugates", Immunol. Rev., 62:119-58 (1982).

Alternatively, an antibody can be conjugated to a second antibody to form an antibody heteroconjugate as described by Segal in U.S. Patent No. 4,676,980.

Accordingly, in one aspect, the invention provides substantially purified antibodies or fragment thereof, and non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103, or an amino acid sequence encoded by the cDNA of a clone deposited as ATCC™ 207176; a fragment of at least 15 amino acid residues of the amino

acid sequence of any one of SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, or 103, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the cDNA of a clone deposited as ATCC™ 207176, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. In various embodiments, the substantially purified antibodies of the invention, or fragments thereof, can be human, non-human, chimeric and/or humanized antibodies.

In another aspect, the invention provides non-human antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited as ATCC™ 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the cDNA of a clone deposited as ATCC™ 207176, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. Such non-human antibodies can be goat, mouse, sheep, horse, chicken, rabbit, or rat antibodies. Alternatively, the non-human antibodies of the invention can be chimeric

and/or humanized antibodies. In addition, the non-human antibodies of the invention can be polyclonal antibodies or monoclonal antibodies.

In still a further aspect, the invention provides monoclonal antibodies or fragments thereof, which antibodies or fragments specifically bind to a polypeptide comprising an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited as ATCC™ 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID Nos:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the cDNA of a clone deposited as any of ATCC™ 207176, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. The monoclonal antibodies can be human, humanized, chimeric and/or non-human antibodies.

The substantially purified antibodies or fragments thereof specifically bind to a signal peptide, a secreted sequence, an extracellular domain, a transmembrane or a cytoplasmic domain cytoplasmic membrane of a polypeptide of the invention. In a particularly preferred embodiment, the substantially purified antibodies or fragments thereof, the non-human antibodies or fragments thereof, and/or the monoclonal antibodies or fragments thereof, of the invention specifically bind to a secreted sequence or an extracellular domain of the amino acid sequence of SEQ ID NOs:26, 27, 31, 33, 35, 39, 41, 46, 47, 51, 122, or 123. Preferably, the secreted sequence or extracellular domain to which the antibody, or fragment thereof, binds comprises from about amino acids 34-488 of SEQ ID NO:2 (SEQ ID NO:26),from amino acids 34-317 of SEQ ID NO:2 (SEQ ID NO:27), from amino acids 29-252 of SEQ ID NO:5 (SEQ ID NO:31), from amino acids 22-226 of SEQ ID NO:8 (SEQ ID NO:33), from amino acids 20-105 of SEQ ID NO:11 (SEQ ID

NO:35) from amino acids 98-317 of SEQ ID NO:2 (SEQ ID NO:41), from amino acids 1-558 of SEQ ID NO:15 (SEQ ID NO:122), from amino acids 773-778 of SEQ ID NO:15 (SEQ ID NO:123), from amino acids 34-79 of SEQ ID NO:17 (SEQ ID NO:46), from amino acids 342-487 of SEQ ID NO:17 (SEQ ID NO:47) from about amino acids 98-317 of SEQ ID NO:17 (SEQ ID NO:51), from amino acids 29-252 of SEQ ID NO: (SEQ ID NO:31), from amino acids 22-226 of SEQ ID NO:16 (SEQ ID NO:33), and amino acids 20-105 of SEQ ID NO:16 (SEQ ID NO:35).

Any of the antibodies of the invention can be conjugated to a therapeutic moiety or to a detectable substance. Non-limiting examples of detectable substances that can be conjugated to the antibodies of the invention are an enzyme, a prosthetic group, a fluorescent material, a luminescent material, a bioluminescent material, and a radioactive material.

The invention also provides a kit containing an antibody of the invention conjugated to a detectable substance, and instructions for use. Still another aspect of the invention is a pharmaceutical composition comprising an antibody of the invention and a pharmaceutically acceptable carrier. In preferred embodiments, the pharmaceutical composition contains an antibody of the invention, a therapeutic moiety, and a pharmaceutically acceptable carrier.

Still another aspect of the invention is a method of making an antibody that specifically recognizes TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267, the method comprising immunizing a mammal with a polypeptide. The polypeptide used as an immungen comprises an amino acid sequence selected from the group consisting of: the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, or an amino acid sequence encoded by the cDNA of a clone deposited as ATCC™ 207176; a fragment of at least 15 amino acid residues of the amino acid sequence of any one of SEQ ID NOs: 2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, an amino acid sequence which is at least 95% identical to the amino acid sequence of any one of SEQ ID NOs:2, 5, 8, 11, 14, 17, 20, 23, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, wherein the percent identity is determined using the ALIGN program of the GCG software package with a PAM120 weight residue table, a gap length penalty of 12, and a gap penalty of 4; and an amino acid sequence which is encoded by a nucleic acid molecule which hybridizes to the nucleic acid molecule consisting of any one of SEQ ID NOs:1, 3, 4, 6, 7,

9, 10, 12, 13, 15, 16, 17, 18, 19, 21, 22, 24, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, or 102, or the cDNA of a clone deposited as ATCC™ 207176, or a complement thereof, under conditions of hybridization of 6X SSC at 45°C and washing in 0.2 X SSC, 0.1% SDS at 65°C. After immunization, a sample is collected from the mammal that contains an antibody that specifically recognizes GPVI. Preferably, the polypeptide is recombinantly produced using a non-human host cell. Optionally, the antibodies can be further purified from the sample using techniques well known to those of skill in the art. The method can further comprise producing a monoclonal antibody-producing cell from the cells of the mammal. Optionally, antibodies are collected from the antibody-producing cell.

III. Recombinant Expression Vectors and Host Cells

Another aspect of the invention pertains to vectors, preferably expression vectors, containing a nucleic acid encoding a polypeptide of the invention (or a portion thereof). As used herein, the term "vector" refers to a nucleic acid molecule capable of transporting another nucleic acid to which it has been linked. One type of vector is a "plasmid", which refers to a circular double stranded DNA loop into which additional DNA segments can be ligated. Another type of vector is a viral vector, wherein additional DNA segments can be ligated into the viral genome. Certain vectors are capable of autonomous replication in a host cell into which they are introduced (e.g., bacterial vectors having a bacterial origin of replication and episomal mammalian vectors). Other vectors (e.g., non-episomal mammalian vectors) are integrated into the genome of a host cell upon introduction into the host cell, and thereby are replicated along with the host genome. Moreover, certain vectors, expression vectors, are capable of directing the expression of genes to which they are operably linked. In general, expression vectors of utility in recombinant DNA techniques are often in the form of plasmids (vectors). However, the invention is intended to include such other forms of expression vectors, such as viral vectors (e.g., replication defective retroviruses, adenoviruses and adeno-associated viruses), which serve equivalent functions.

The recombinant expression vectors of the invention comprise a nucleic acid of the invention in a form suitable for expression of the nucleic acid in a host cell. This means that the recombinant expression vectors include one or more regulatory sequences, selected on the basis of the host cells to be used for expression, which is operably linked to the nucleic acid sequence to be expressed. Within a recombinant expression vector, "operably linked" is intended to mean that the nucleotide sequence of interest is linked to the regulatory

sequence(s) in a manner which allows for expression of the nucleotide sequence (e.g., in an *in vitro* transcription/translation system or in a host cell when the vector is introduced into the host cell). The term "regulatory sequence" is intended to include promoters, enhancers and other expression control elements (e.g., polyadenylation signals). Such regulatory sequences are described, for example, in Goeddel, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, CA (1990). Regulatory sequences include those which direct constitutive expression of a nucleotide sequence in many types of host cell and those which direct expression of the nucleotide sequence only in certain host cells (e.g., tissue-specific regulatory sequences). It will be appreciated by those skilled in the art that the design of the expression vector can depend on such factors as the choice of the host cell to be transformed, the level of expression of protein desired, etc. The expression vectors of the invention can be introduced into host cells to thereby produce proteins or peptides, including fusion proteins or peptides, encoded by nucleic acids as described herein.

The recombinant expression vectors of the invention can be designed for expression of a polypeptide of the invention in prokaryotic (e.g., *E. coli*) or eukaryotic cells (e.g., insect cells (using baculovirus expression vectors), yeast cells or mammalian cells). Suitable host cells are discussed further in Goeddel, *supra*. Alternatively, the recombinant expression vector can be transcribed and translated *in vitro*, for example using T7 promoter regulatory sequences and T7 polymerase.

Expression of proteins in prokaryotes is most often carried out in *E. coli* with vectors containing constitutive or inducible promoters directing the expression of either fusion or non-fusion proteins. Fusion vectors add a number of amino acids to a protein encoded therein, usually to the amino terminus of the recombinant protein. Such fusion vectors typically serve three purposes: 1) to increase expression of recombinant protein; 2) to increase the solubility of the recombinant protein; and 3) to aid in the purification of the recombinant protein by acting as a ligand in affinity purification. Often, in fusion expression vectors, a proteolytic cleavage site is introduced at the junction of the fusion moiety and the recombinant protein to enable separation of the recombinant protein from the fusion moiety subsequent to purification of the fusion protein. Such enzymes, and their cognate recognition sequences, include Factor Xa, thrombin and enterokinase. Typical fusion expression vectors include pGEX (Pharmacia Biotech Inc; Smith and Johnson (1988) *Gene* 67:31-40), pMAL (New England Biolabs, Beverly, MA) and pRIT5 (Pharmacia, Piscataway, NJ) which fuse glutathione S-transferase (GST), maltose E binding protein, or protein A, respectively, to the target recombinant protein.

Examples of suitable inducible non-fusion *E. coli* expression vectors include pTrc (Amann et al., (1988) *Gene* 69:301-315) and pET 11d (Studier et al., *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 60-89). Target gene expression from the pTrc vector relies on host RNA polymerase transcription from a hybrid trp-lac fusion promoter. Target gene expression from the pET 11d vector relies on transcription from a T7 gn10-lac fusion promoter mediated by a coexpressed viral RNA polymerase (T7 gn1). This viral polymerase is supplied by host strains BL21(DE3) or HMS174(DE3) from a resident ë prophage harboring a T7 gn1 gene under the transcriptional control of the lacUV 5 promoter.

One strategy to maximize recombinant protein expression in *E. coli* is to express the protein in a host bacteria with an impaired capacity to proteolytically cleave the recombinant protein (Gottesman, *Gene Expression Technology: Methods in Enzymology* 185, Academic Press, San Diego, California (1990) 119-128). Another strategy is to alter the nucleic acid sequence of the nucleic acid to be inserted into an expression vector so that the individual codons for each amino acid are those preferentially utilized in *E. coli* (Wada et al. (1992) *Nucleic Acids Res.* 20:2111-2118). Such alteration of nucleic acid sequences of the invention can be carried out by standard DNA synthesis techniques.

In another embodiment, the expression vector is a yeast expression vector. Examples of vectors for expression in yeast *S. cerivisae* include pYepSec1 (Baldari et al. (1987) *EMBO J.* 6:229-234), pMFa (Kurjan and Herskowitz, (1982) *Cell* 30:933-943), pJRY88 (Schultz et al. (1987) *Gene* 54:113-123), pYES2 (Invitrogen Corporation, San Diego, CA), and pPicZ (Invitrogen Corp, San Diego, CA).

Alternatively, the expression vector is a baculovirus expression vector. Baculovirus vectors available for expression of proteins in cultured insect cells (e.g., Sf 9 cells) include the pAc series (Smith et al. (1983) *Mol. Cell Biol.* 3:2156-2165) and the pVL series (Lucklow and Summers (1989) *Virology* 170:31-39).

In yet another embodiment, a nucleic acid of the invention is expressed in mammalian cells using a mammalian expression vector. Examples of mammalian expression vectors include pCDM8 (Seed (1987) *Nature* 329:840) and pMT2PC (Kaufman et al. (1987) *EMBO J.* 6:187-195). When used in mammalian cells, the expression vector's control functions are often provided by viral regulatory elements. For example, commonly used promoters are derived from polyoma, Adenovirus 2, cytomegalovirus and Simian Virus 40. For other suitable expression systems for both prokaryotic and eukaryotic cells see chapters 16 and 17 of Sambrook et al., *supra*.

In another embodiment, the recombinant mammalian expression vector is capable of directing expression of the nucleic acid preferentially in a particular cell type (e.g., tissue-specific regulatory elements are used to express the nucleic acid). Tissue-specific regulatory elements are known in the art. Non-limiting examples of suitable tissue-specific promoters include the albumin promoter (liver-specific; Pinkert et al. (1987) *Genes Dev.* 1:268-277), lymphoid-specific promoters (Calame and Eaton (1988) *Adv. Immunol.* 43:235-275), in particular promoters of T cell receptors (Winoto and Baltimore (1989) *EMBO J.* 8:729-733) and immunoglobulins (Banerji et al. (1983) *Cell* 33:729-740; Queen and Baltimore (1983) *Cell* 33:741-748), neuron-specific promoters (e.g., the neurofilament promoter; Byrne and Ruddle (1989) *Proc. Natl. Acad. Sci. USA* 86:5473-5477), pancreas-specific promoters (Edlund et al. (1985) *Science* 230:912-916), and mammary gland-specific promoters (e.g., milk whey promoter; U.S. Patent No. 4,873,316 and European Application Publication No. 264,166). Developmentally-regulated promoters are also encompassed, for example the murine hox promoters (Kessel and Gruss (1990) *Science* 249:374-379) and the α-fetoprotein promoter (Campes and Tilghman (1989) *Genes Dev.* 3:537-546).

The invention further provides a recombinant expression vector comprising a DNA molecule of the invention cloned into the expression vector in an antisense orientation. That is, the DNA molecule is operably linked to a regulatory sequence in a manner which allows for expression (by transcription of the DNA molecule) of an RNA molecule which is antisense to the mRNA encoding a polypeptide of the invention. Regulatory sequences operably linked to a nucleic acid cloned in the antisense orientation can be chosen which direct the continuous expression of the antisense RNA molecule in a variety of cell types, for instance viral promoters and/or enhancers, or regulatory sequences can be chosen which direct constitutive, tissue specific or cell type specific expression of antisense RNA. The antisense expression vector can be in the form of a recombinant plasmid, phagemid or attenuated virus in which antisense nucleic acids are produced under the control of a high efficiency regulatory region, the activity of which can be determined by the cell type into which the vector is introduced. For a discussion of the regulation of gene expression using antisense genes see Weintraub et al. (*Reviews - Trends in Genetics*, Vol. 1(1) 1986).

Another aspect of the invention pertains to host cells into which a recombinant expression vector of the invention has been introduced. The terms "host cell" and "recombinant host cell" are used interchangeably herein. It is understood that such terms refer not only to the particular subject cell but to the progeny or potential progeny of such a cell. Because certain modifications may occur in succeeding generations due to either

mutation or environmental influences, such progeny may not, in fact, be identical to the parent cell, but are still included within the scope of the term as used herein.

A host cell can be any prokaryotic (e.g., *E. coli*) or eukaryotic cell (e.g., insect cells, yeast or mammalian cells).

Vector DNA can be introduced into prokaryotic or eukaryotic cells via conventional transformation or transfection techniques. As used herein, the terms "transformation" and "transfection" are intended to refer to a variety of art-recognized techniques for introducing foreign nucleic acid into a host cell, including calcium phosphate or calcium chloride co-precipitation, DEAE-dextran-mediated transfection, lipofection, or electroporation. Suitable methods for transforming or transfecting host cells can be found in Sambrook, et al. (*supra*), and other laboratory manuals.

For stable transfection of mammalian cells, it is known that, depending upon the expression vector and transfection technique used, only a small fraction of cells may integrate the foreign DNA into their genome. In order to identify and select these integrants, a gene that encodes a selectable marker (e.g., for resistance to antibiotics) is generally introduced into the host cells along with the gene of interest. Preferred selectable markers include those which confer resistance to drugs, such as G418, hygromycin and methotrexate. Cells stably transfected with the introduced nucleic acid can be identified by drug selection (e.g., cells that have incorporated the selectable marker gene will survive, while the other cells die).

In another embodiment, the expression characteristics of an endogenous (*e.g.*, TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267) nucleic acid within a cell, cell line or microorganism may be modified by inserting a DNA regulatory element heterologous to the endogenous gene of interest into the genome of a cell, stable cell line or cloned microorganism such that the inserted regulatory element is operatively linked with the endogenous gene (*e.g.*, TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267) and controls, modulates or activates the endogenous gene. For example, endogenous TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 which are normally "transcriptionally silent", *i.e.*, TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 genes which are normally not expressed, or are expressed only at very low levels in a cell line or microorganism, may be activated by inserting a regulatory element which is capable of promoting the expression of a normally expressed gene product in that cell line or microorganism. Alternatively, transcriptionally silent, endogenous TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO

267 genes may be activated by insertion of a promiscuous regulatory element that works across cell types.

A heterologous regulatory element may be inserted into a stable cell line or cloned microorganism, such that it is operatively linked with and activates expression of endogenous TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 genes, using techniques, such as targeted homologous recombination, which are well known to those of skill in the art, and described *e.g.*, in Chappel, U.S. Patent No. 5,272,071; PCT publication No. WO 91/06667, published May 16, 1991.

A host cell of the invention, such as a prokaryotic or eukaryotic host cell in culture, can be used to produce a polypeptide of the invention. Accordingly, the invention further provides methods for producing a polypeptide of the invention using the host cells of the invention. In one embodiment, the method comprises culturing the host cell of invention (into which a recombinant expression vector encoding a polypeptide of the invention has been introduced) in a suitable medium such that the polypeptide is produced. In another embodiment, the method further comprises isolating the polypeptide from the medium or the host cell.

The host cells of the invention can also be used to produce nonhuman transgenic animals. For example, in one embodiment, a host cell of the invention is a fertilized oocyte or an embryonic stem cell into which a sequences encoding a polypeptide of the invention have been introduced. Such host cells can then be used to create non-human transgenic animals in which exogenous sequences encoding a polypeptide of the invention have been introduced into their genome or homologous recombinant animals in which endogenous encoding a polypeptide of the invention sequences have been altered. Such animals are useful for studying the function and/or activity of the polypeptide and for identifying and/or evaluating modulators of polypeptide activity. As used herein, a "transgenic animal" is a non-human animal, preferably a mammal, more preferably a rodent such as a rat or mouse, in which one or more of the cells of the animal includes a transgene. Other examples of transgenic animals include non-human primates, sheep, dogs, cows, goats, chickens, amphibians, etc. A transgene is exogenous DNA which is integrated into the genome of a cell from which a transgenic animal develops and which remains in the genome of the mature animal, thereby directing the expression of an encoded gene product in one or more cell types or tissues of the transgenic animal. As used herein, an "homologous recombinant animal" is a non-human animal, preferably a mammal, more preferably a mouse, in which an endogenous gene has been altered by homologous recombination between the

endogenous gene and an exogenous DNA molecule introduced into a cell of the animal, e.g., an embryonic cell of the animal, prior to development of the animal.

A transgenic animal of the invention can be created by introducing nucleic acid encoding a polypeptide of the invention (or a homologue thereof) into the male pronuclei of a fertilized oocyte, e.g., by microinjection, retroviral infection, and allowing the oocyte to develop in a pseudopregnant female foster animal. Intronic sequences and polyadenylation signals can also be included in the transgene to increase the efficiency of expression of the transgene. A tissue-specific regulatory sequence(s) can be operably linked to the transgene to direct expression of the polypeptide of the invention to particular cells. Methods for generating transgenic animals via embryo manipulation and microinjection, particularly animals such as mice, have become conventional in the art and are described, for example, in U.S. Patent NOS. 4,736,866 and 4,870,009, U.S. Patent No. 4,873,191 and in Hogan, *Manipulating the Mouse Embryo*, (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 1986). Similar methods are used for production of other transgenic animals. A transgenic founder animal can be identified based upon the presence of the transgene in its genome and/or expression of mRNA encoding the transgene in tissues or cells of the animals. A transgenic founder animal can then be used to breed additional animals carrying the transgene. Moreover, transgenic animals carrying the transgene can further be bred to other transgenic animals carrying other transgenes.

To create an homologous recombinant animal, a vector is prepared which contains at least a portion of a gene encoding a polypeptide of the invention into which a deletion, addition or substitution has been introduced to thereby alter, e.g., functionally disrupt, the gene. In a preferred embodiment, the vector is designed such that, upon homologous recombination, the endogenous gene is functionally disrupted (i.e., no longer encodes a functional protein; also referred to as a "knock out" vector). Alternatively, the vector can be designed such that, upon homologous recombination, the endogenous gene is mutated or otherwise altered but still encodes functional protein (e.g., the upstream regulatory region can be altered to thereby alter the expression of the endogenous protein). In the homologous recombination vector, the altered portion of the gene is flanked at its 5' and 3' ends by additional nucleic acid of the gene to allow for homologous recombination to occur between the exogenous gene carried by the vector and an endogenous gene in an embryonic stem cell. The additional flanking nucleic acid sequences are of sufficient length for successful homologous recombination with the endogenous gene. Typically, several kilobases of flanking DNA (both at the 5' and 3' ends) are included in the vector (*see, e.g.*, Thomas and Capecchi (1987) *Cell* 51:503 for a description of homologous recombination vectors). The

vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced gene has homologously recombined with the endogenous gene are selected (*see, e.g.,* Li et al. (1992) *Cell* 69:915). The selected cells are then injected into a blastocyst of an animal (e.g., a mouse) to form aggregation chimeras (*see, e.g.,* Bradley in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, Robertson, ed. (IRL, Oxford, 1987) pp. 113-152). A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term. Progeny harboring the homologously recombined DNA in their germ cells can be used to breed animals in which all cells of the animal contain the homologously recombined DNA by germline transmission of the transgene. Methods for constructing homologous recombination vectors and homologous recombinant animals are described further in Bradley (1991) *Current Opinion in Bio/Technology* 2:823-829 and in PCT Publication NOS. WO 90/11354, WO 91/01140, WO 92/0968, and WO 93/04169.

In another embodiment, transgenic non-human animals can be produced which contain selected systems which allow for regulated expression of the transgene. One example of such a system is the *cre/loxP* recombinase system of bacteriophage P1. For a description of the *cre/loxP* recombinase system, *see, e.g.,* Lakso et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:6232-6236. Another example of a recombinase system is the FLP recombinase system of *Saccharomyces cerevisiae* (O'Gorman et al. (1991) *Science* 251:1351-1355. If a *cre/loxP* recombinase system is used to regulate expression of the transgene, animals containing transgenes encoding both the *Cre* recombinase and a selected protein are required. Such animals can be provided through the construction of "double" transgenic animals, e.g., by mating two transgenic animals, one containing a transgene encoding a selected protein and the other containing a transgene encoding a recombinase.

Clones of the non-human transgenic animals described herein can also be produced according to the methods described in Wilmut et al. (1997) *Nature* 385:810-813 and PCT Publication NOS. WO 97/07668 and WO 97/07669.

IV. Pharmaceutical Compositions

The nucleic acid molecules, polypeptides, and antibodies (also referred to herein as "active compounds") of the invention can be incorporated into pharmaceutical compositions suitable for administration. Such compositions typically comprise the nucleic acid molecule, protein, or antibody and a pharmaceutically acceptable carrier. As used herein the language "pharmaceutically acceptable carrier" is intended to include any and all solvents,

dispersion media, coatings, antibacterial and antifungal agents, isotonic and absorption delaying agents, and the like, compatible with pharmaceutical administration. The use of such media and agents for pharmaceutically active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active compound, use thereof in the compositions is contemplated. Supplementary active compounds can also be incorporated into the compositions.

The invention includes methods for preparing pharmaceutical compositions for modulating the expression or activity of a polypeptide or nucleic acid of the invention. Such methods comprise formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention. Such compositions can further include additional active agents. Thus, the invention further includes methods for preparing a pharmaceutical composition by formulating a pharmaceutically acceptable carrier with an agent which modulates expression or activity of a polypeptide or nucleic acid of the invention and one or more addtional active compounds.

A pharmaceutical composition of the invention is formulated to be compatible with its intended route of administration. Examples of routes of administration include parenteral, e.g., intravenous, intradermal, subcutaneous, oral (e.g., inhalation), transdermal (topical), transmucosal, and rectal administration. Solutions or suspensions used for parenteral, intradermal, or subcutaneous application can include the following components: a sterile diluent such as water for injection, saline solution, fixed oils, polyethylene glycols, glycerine, propylene glycol or other synthetic solvents; antibacterial agents such as benzyl alcohol or methyl parabens; antioxidants such as ascorbic acid or sodium bisulfite; chelating agents such as ethylenediaminetetraacetic acid; buffers such as acetates, citrates or phosphates and agents for the adjustment of tonicity such as sodium chloride or dextrose. pH can be adjusted with acids or bases, such as hydrochloric acid or sodium hydroxide. The parenteral preparation can be enclosed in ampoules, disposable syringes or multiple dose vials made of glass or plastic.

Pharmaceutical compositions suitable for injectable use include sterile aqueous solutions (where water soluble) or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions. For intravenous administration, suitable carriers include physiological saline, bacteriostatic water, Cremophor EL (BASF; Parsippany, NJ) or phosphate buffered saline (PBS). In all cases, the composition must be sterile and should be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating

action of microorganisms such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (for example, glycerol, propylene glycol, and liquid polyetheylene glycol, and the like), and suitable mixtures thereof. The proper fluidity can be maintained, for example, by the use of a coating such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. Prevention of the action of microorganisms can be achieved by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, ascorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars, polyalcohols such as mannitol, sorbitol, sodium chloride in the composition. Prolonged absorption of the injectable compositions can be brought about by including in the composition an agent which delays absorption, for example, aluminum monostearate and gelatin.

Sterile injectable solutions can be prepared by incorporating the active compound (e.g., a polypeptide or antibody) in the required amount in an appropriate solvent with one or a combination of ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the active compound into a sterile vehicle which contains a basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum drying and freeze-drying which yields a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

Oral compositions generally include an inert diluent or an edible carrier. They can be enclosed in gelatin capsules or compressed into tablets. For the purpose of oral therapeutic administration, the active compound can be incorporated with excipients and used in the form of tablets, troches, or capsules. Oral compositions can also be prepared using a fluid carrier for use as a mouthwash, wherein the compound in the fluid carrier is applied orally and swished and expectorated or swallowed.

Pharmaceutically compatible binding agents, and/or adjuvant materials can be included as part of the composition. The tablets, pills, capsules, troches and the like can contain any of the following ingredients, or compounds of a similar nature: a binder such as microcrystalline cellulose, gum tragacanth or gelatin; an excipient such as starch or lactose, a disintegrating agent such as alginic acid, Primogel, or corn starch; a lubricant such as magnesium stearate or Sterotes; a glidant such as colloidal silicon dioxide; a sweetening

agent such as sucrose or saccharin; or a flavoring agent such as peppermint, methyl salicylate, or orange flavoring.

For administration by inhalation, the compounds are delivered in the form of an aerosol spray from a pressurized container or dispenser which contains a suitable propellant, e.g., a gas such as carbon dioxide, or a nebulizer.

Systemic administration can also be by transmucosal or transdermal means. For transmucosal or transdermal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art, and include, for example, for transmucosal administration, detergents, bile salts, and fusidic acid derivatives. Transmucosal administration can be accomplished through the use of nasal sprays or suppositories. For transdermal administration, the active compounds are formulated into ointments, salves, gels, or creams as generally known in the art.

The compounds can also be prepared in the form of suppositories (e.g., with conventional suppository bases such as cocoa butter and other glycerides) or retention enemas for rectal delivery.

In one embodiment, the active compounds are prepared with carriers that will protect the compound against rapid elimination from the body, such as a controlled release formulation, including implants and microencapsulated delivery systems. Biodegradable, biocompatible polymers can be used, such as ethylene vinyl acetate, polyanhydrides, polyglycolic acid, collagen, polyorthoesters, and polylactic acid. Methods for preparation of such formulations will be apparent to those skilled in the art. The materials can also be obtained commercially from Alza Corporation and Nova Pharmaceuticals, Inc. Liposomal suspensions (including liposomes targeted to infected cells with monoclonal antibodies to viral antigens) can also be used as pharmaceutically acceptable carriers. These can be prepared according to methods known to those skilled in the art, for example, as described in U.S. Patent No. 4,522,811.

It is especially advantageous to formulate oral or parenteral compositions in dosage unit form for ease of administration and uniformity of dosage. Dosage unit form as used herein refers to physically discrete units suited as unitary dosages for the subject to be treated; each unit containing a predetermined quantity of active compound calculated to produce the desired therapeutic effect in association with the required pharmaceutical carrier. The specification for the dosage unit forms of the invention are dictated by and directly dependent on the unique characteristics of the active compound and the particular

therapeutic effect to be achieved, and the limitations inherent in the art of compounding such an active compound for the treatment of individuals.

For antibodies, the preferred dosage is 0.1 mg/kg to 100 mg/kg of body weight (generally 10 mg/kg to 20 mg/kg). If the antibody is to act in the brain, a dosage of 50 mg/kg to 100 mg/kg is usually appropriate. Generally, partially human antibodies and fully human antibodies have a longer half-life within the human body than other antibodies. Accordingly, lower dosages and less frequent administration is often possible. Modifications such as lipidation can be used to stabilize antibodies and to enhance uptake and tissue penetration (e.g., into the brain). A method for lipidation of antibodies is described by Cruikshank et al. ((1997) *J. Acquired Immune Deficiency Syndromes and Human Retrovirology* 14:193).

The nucleic acid molecules of the invention can be inserted into vectors and used as gene therapy vectors. Gene therapy vectors can be delivered to a subject by, for example, intravenous injection, local administration (U.S. Patent 5,328,470) or by stereotactic injection (*see, e.g.,* Chen et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:3054-3057). The pharmaceutical preparation of the gene therapy vector can include the gene therapy vector in an acceptable diluent, or can comprise a slow release matrix in which the gene delivery vehicle is imbedded. Alternatively, where the complete gene delivery vector can be produced intact from recombinant cells, e.g. retroviral vectors, the pharmaceutical preparation can include one or more cells which produce the gene delivery system.

The pharmaceutical compositions can be included in a container, pack, or dispenser together with instructions for administration.

V. Uses and Methods of the Invention

The nucleic acid molecules, proteins, protein homologues, and antibodies described herein can be used in one or more of the following methods: a) screening assays; b) detection assays (e.g., chromosomal mapping, tissue typing, forensic biology); c) predictive medicine (e.g., diagnostic assays, prognostic assays, monitoring clinical trials, and pharmacogenomics); and d) methods of treatment (e.g., therapeutic and prophylactic). For example, polypeptides of the invention can to used to (i) modulate cellular proliferation; (ii) modulate cellular differentiation; and/or (iii) modulate cellular adhesion. The isolated nucleic acid molecules of the invention can be used to express proteins (e.g., via a recombinant expression vector in a host cell in gene therapy applications), to detect mRNA

(e.g., in a biological sample) or a genetic lesion, and to modulate activity of a polypeptide of the invention. In addition, the polypeptides of the invention can be used to screen drugs or compounds which modulate activity or expression of a polypeptide of the invention as well as to treat disorders characterized by insufficient or excessive production of a protein of the invention or production of a form of a protein of the invention which has decreased or aberrant activity compared to the wild type protein. In addition, the antibodies of the invention can be used to detect and isolate a protein of the and modulate activity of a protein of the invention.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

A. Screening Assays

The invention provides a method (also referred to herein as a "screening assay") for identifying modulators, i.e., candidate or test compounds or agents (e.g., peptides, peptidomimetics, small molecules or other drugs) which bind to polypeptide of the invention or have a stimulatory or inhibitory effect on, for example, expression or activity of a polypeptide of the invention.

In one embodiment, the invention provides assays for screening candidate or test compounds which bind to or modulate the activity of the membrane-bound form of a polypeptide of the invention or biologically active portion thereof. The test compounds of the present invention can be obtained using any of the numerous approaches in combinatorial library methods known in the art, including: biological libraries; spatially addressable parallel solid phase or solution phase libraries; synthetic library methods requiring deconvolution; the "one-bead one-compound" library method; and synthetic library methods using affinity chromatography selection. The biological library approach is limited to peptide libraries, while the other four approaches are applicable to peptide, non-peptide oligomer or small molecule libraries of compounds (Lam (1997) *Anticancer Drug Des.* 12:145).

Examples of methods for the synthesis of molecular libraries can be found in the art, for example in: DeWitt et al. (1993) *Proc. Natl. Acad. Sci. USA* 90:6909; Erb et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:11422; Zuckermann et al. (1994). *J. Med. Chem.* 37:2678; Cho et al. (1993) *Science* 261:1303; Carrell et al. (1994) *Angew. Chem. Int. Ed. Engl.*

33:2059; Carell et al. (1994) *Angew. Chem. Int. Ed. Engl.* 33:2061; and Gallop et al. (1994) *J. Med. Chem.* 37:1233.

Libraries of compounds may be presented in solution (e.g., Houghten (1992) *Bio/Techniques* 13:412-421), or on beads (Lam (1991) *Nature* 354:82-84), chips (Fodor (1993) *Nature* 364:555-556), bacteria (U.S. Patent No. 5,223,409), spores (Patent NOS. 5,571,698; 5,403,484; and 5,223,409), plasmids (Cull et al. (1992) *Proc. Natl. Acad. Sci. USA* 89:1865-1869) or phage (Scott and Smith (1990) *Science* 249:386-390; Devlin (1990) *Science* 249:404-406; Cwirla et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382; and Felici (1991) *J. Mol. Biol.* 222:301-310).

In one embodiment, an assay is a cell-based assay in which a cell which expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface is contacted with a test compound and the ability of the test compound to bind to the polypeptide determined. The cell, for example, can be a yeast cell or a cell of mammalian origin. Determining the ability of the test compound to bind to the polypeptide can be accomplished, for example, by coupling the test compound with a radioisotope or enzymatic label such that binding of the test compound to the polypeptide or biologically active portion thereof can be determined by detecting the labeled compound in a complex. For example, test compounds can be labeled with ^{125}I, ^{35}S, ^{14}C, or ^{3}H, either directly or indirectly, and the radioisotope detected by direct counting of radioemmission or by scintillation counting. Alternatively, test compounds can be enzymatically labeled with, for example, horseradish peroxidase, alkaline phosphatase, or luciferase, and the enzymatic label detected by determination of conversion of an appropriate substrate to product. In a preferred embodiment, the assay comprises contacting a cell which expresses a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind to the polypeptide or a biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-based assay comprising contacting a cell expressing a membrane-bound form of a polypeptide of the invention, or a biologically active portion thereof, on the cell surface with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide

or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of the polypeptide or a biologically active portion thereof can be accomplished, for example, by determining the ability of the polypeptide protein to bind to or interact with a target molecule.

Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by one of the methods described above for determining direct binding. As used herein, a "target molecule" is a molecule with which a selected polypeptide (e.g., a polypeptide of the invention binds or interacts with in nature, for example, a molecule on the surface of a cell which expresses the selected protein, a molecule on the surface of a second cell, a molecule in the extracellular milieu, a molecule associated with the internal surface of a cell membrane or a cytoplasmic molecule. A target molecule can be a polypeptide of the invention or some other polypeptide or protein. For example, a target molecule can be a component of a signal transduction pathway which facilitates transduction of an extracellular signal (e.g., a signal generated by binding of a compound to a polypeptide of the invention) through the cell membrane and into the cell or a second intercellular protein which has catalytic activity or a protein which facilitates the association of downstream signaling molecules with a polypeptide of the invention. Determining the ability of a polypeptide of the invention to bind to or interact with a target molecule can be accomplished by determining the activity of the target molecule. For example, the activity of the target molecule can be determined by detecting induction of a cellular second messenger of the target (e.g., intracellular Ca^{2+}, diacylglycerol, IP3, etc.), detecting catalytic/enzymatic activity of the target on an appropriate substrate, detecting the induction of a reporter gene (e.g., a regulatory element that is responsive to a polypeptide of the invention operably linked to a nucleic acid encoding a detectable marker, e.g. luciferase), or detecting a cellular response, for example, cellular differentiation, or cell proliferation.

In yet another embodiment, an assay of the present invention is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the ability of the test compound to bind to the polypeptide or biologically active portion thereof. Binding of the test compound to the polypeptide can be determined either directly or indirectly as described above. In a preferred embodiment, the assay includes contacting the polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the

ability of the test compound to interact with the polypeptide comprises determining the ability of the test compound to preferentially bind to the polypeptide or biologically active portion thereof as compared to the known compound.

In another embodiment, an assay is a cell-free assay comprising contacting a polypeptide of the invention or biologically active portion thereof with a test compound and determining the ability of the test compound to modulate (e.g., stimulate or inhibit) the activity of the polypeptide or biologically active portion thereof. Determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished, for example, by determining the ability of the polypeptide to bind to a target molecule by one of the methods described above for determining direct binding. In an alternative embodiment, determining the ability of the test compound to modulate the activity of the polypeptide can be accomplished by determining the ability of the polypeptide of the invention to further modulate the target molecule. For example, the catalytic/enzymatic activity of the target molecule on an appropriate substrate can be determined as previously described.

In yet another embodiment, the cell-free assay comprises contacting a polypeptide of the invention or biologically active portion thereof with a known compound which binds the polypeptide to form an assay mixture, contacting the assay mixture with a test compound, and determining the ability of the test compound to interact with the polypeptide, wherein determining the ability of the test compound to interact with the polypeptide comprises determining the ability of the polypeptide to preferentially bind to or modulate the activity of a target molecule.

The cell-free assays of the present invention are amenable to use of both a soluble form or the membrane-bound form of a polypeptide of the invention. In the case of cell-free assays comprising the membrane-bound form of the polypeptide, it may be desirable to utilize a solubilizing agent such that the membrane-bound form of the polypeptide is maintained in solution. Examples of such solubilizing agents include non-ionic detergents such as n-octylglucoside, n-dodecylglucoside, n-octylmaltoside, octanoyl-N-methylglucamide, decanoyl-N-methylglucamide, Triton X-100, Triton X-114, Thesit, Isotridecypoly(ethylene glycol ether)n, 3-[(3-cholamidopropyl)dimethylamminio]-1-propane sulfonate (CHAPS), 3-[(3-cholamidopropyl)dimethylamminio]-2-hydroxy-1-propane sulfonate (CHAPSO), or N-dodecyl=N,N-dimethyl-3-ammonio-1-propane sulfonate.

In more than one embodiment of the above assay methods of the present invention, it may be desirable to immobilize either the polypeptide of the invention or its target molecule to facilitate separation of complexed from uncomplexed forms of one or both of the

proteins, as well as to accommodate automation of the assay. Binding of a test compound to the polypeptide, or interaction of the polypeptide with a target molecule in the presence and absence of a candidate compound, can be accomplished in any vessel suitable for containing the reactants. Examples of such vessels include microtitre plates, test tubes, and micro-centrifuge tubes. In one embodiment, a fusion protein can be provided which adds a domain that allows one or both of the proteins to be bound to a matrix. For example, glutathione-S-transferase fusion proteins or glutathione-S-transferase fusion proteins can be adsorbed onto glutathione sepharose beads (Sigma Chemical; St. Louis, MO) or glutathione derivatized microtitre plates, which are then combined with the test compound or the test compound and either the non-adsorbed target protein or A polypeptide of the invention, and the mixture incubated under conditions conducive to complex formation (e.g., at physiological conditions for salt and pH). Following incubation, the beads or microtitre plate wells are washed to remove any unbound components and complex formation is measured either directly or indirectly, for example, as described above. Alternatively, the complexes can be dissociated from the matrix, and the level of binding or activity of the polypeptide of the invention can be determined using standard techniques.

Other techniques for immobilizing proteins on matrices can also be used in the screening assays of the invention. For example, either the polypeptide of the invention or its target molecule can be immobilized utilizing conjugation of biotin and streptavidin. Biotinylated polypeptide of the invention or target molecules can be prepared from biotin-NHS (N-hydroxy-succinimide) using techniques well known in the art (e.g., biotinylation kit, Pierce Chemicals; Rockford, IL), and immobilized in the wells of streptavidin-coated 96 well plates (Pierce Chemical). Alternatively, antibodies reactive with the polypeptide of the invention or target molecules but which do not interfere with binding of the polypeptide of the invention to its target molecule can be derivatized to the wells of the plate, and unbound target or polypeptidede of the invention trapped in the wells by antibody conjugation. Methods for detecting such complexes, in addition to those described above for the GST-immobilized complexes, include immunodetection of complexes using antibodies reactive with the polypeptide of the invention or target molecule, as well as enzyme-linked assays which rely on detecting an enzymatic activity associated with the polypeptide of the invention or target molecule.

In another embodiment, modulators of expression of a polypeptide of the invention are identified in a method in which a cell is contacted with a candidate compound and the expression of the selected mRNA or protein (i.e., the mRNA or protein corresponding to a polypeptide or nucleic acid of the invention) in the cell is determined. The level of

expression of the selected mRNA or protein in the presence of the candidate compound is compared to the level of expression of the selected mRNA or protein in the absence of the candidate compound. The candidate compound can then be identified as a modulator of expression of the polypeptide of the invention based on this comparison. For example, when expression of the selected mRNA or protein is greater (statistically significantly greater) in the presence of the candidate compound than in its absence, the candidate compound is identified as a stimulator of the selected mRNA or protein expression. Alternatively, when expression of the selected mRNA or protein is less (statistically significantly less) in the presence of the candidate compound than in its absence, the candidate compound is identified as an inhibitor of the selected mRNA or protein expression. The level of the selected mRNA or protein expression in the cells can be determined by methods described herein.

In yet another aspect of the invention, a polypeptide of the inventions can be used as "bait proteins" in a two-hybrid assay or three hybrid assay (*see, e.g.*, U.S. Patent No. 5,283,317; Zervos et al. (1993) *Cell* 72:223-232; Madura et al. (1993) *J. Biol. Chem.* 268:12046-12054; Bartel et al. (1993) *Bio/Techniques* 14:920-924; Iwabuchi et al. (1993) *Oncogene* 8:1693-1696; and PCT Publication No. WO 94/10300), to identify other proteins, which bind to or interact with the polypeptide of the invention and modulate activity of the polypeptide of the invention. Such binding proteins are also likely to be involved in the propagation of signals by the polypeptide of the inventions as, for example, upstream or downstream elements of a signaling pathway involving the polypeptide of the invention.

This invention further pertains to novel agents identified by the above-described screening assays and uses thereof for treatments as described herein.

B. Detection Assays

Portions or fragments of the cDNA sequences identified herein (and the corresponding complete gene sequences) can be used in numerous ways as polynucleotide reagents. For example, these sequences can be used to: (i) map their respective genes on a chromosome and, thus, locate gene regions associated with genetic disease; (ii) identify an individual from a minute biological sample (tissue typing); and (iii) aid in forensic identification of a biological sample. These applications are described in the subsections below.

1. Chromosome Mapping

Once the sequence (or a portion of the sequence) of a gene has been isolated, this sequence can be used to map the location of the gene on a chromosome. Accordingly, nucleic acid molecules described herein or fragments thereof, can be used to map the location of the corresponding genes on a chromosome. The mapping of the sequences to chromosomes is an important first step in correlating these sequences with genes associated with disease.

Briefly, genes can be mapped to chromosomes by preparing PCR primers (preferably 15-25 bp in length) from the sequence of a gene of the invention. Computer analysis of the sequence of a gene of the invention can be used to rapidly select primers that do not span more than one exon in the genomic DNA, thus complicating the amplification process. These primers can then be used for PCR screening of somatic cell hybrids containing individual human chromosomes. Only those hybrids containing the human gene corresponding to the gene sequences will yield an amplified fragment. For a review of this technique, see D'Eustachio et al. ((1983) *Science* 220:919-924).

PCR mapping of somatic cell hybrids is a rapid procedure for assigning a particular sequence to a particular chromosome. Three or more sequences can be assigned per day using a single thermal cycler. Using the nucleic acid sequences of the invention to design oligonucleotide primers, sublocalization can be achieved with panels of fragments from specific chromosomes. Other mapping strategies which can similarly be used to map a gene to its chromosome include *in situ* hybridization (described in Fan et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:6223-27), pre-screening with labeled flow-sorted chromosomes, and pre-selection by hybridization to chromosome specific cDNA libraries. Fluorescence *in situ* hybridization (FISH) of a DNA sequence to a metaphase chromosomal spread can further be used to provide a precise chromosomal location in one step. For a review of this technique, see Verma et al. (Human Chromosomes: A Manual of Basic Techniques (Pergamon Press, New York, 1988)).

Reagents for chromosome mapping can be used individually to mark a single chromosome or a single site on that chromosome, or panels of reagents can be used for marking multiple sites and/or multiple chromosomes. Reagents corresponding to noncoding regions of the genes actually are preferred for mapping purposes. Coding sequences are more likely to be conserved within gene families, thus increasing the chance of cross hybridizations during chromosomal mapping.

Once a sequence has been mapped to a precise chromosomal location, the physical position of the sequence on the chromosome can be correlated with genetic map data. (Such data are found, for example, in V. McKusick, Mendelian Inheritance in Man, available on-line through Johns Hopkins University Welch Medical Library). The relationship between genes and disease, mapped to the same chromosomal region, can then be identified through linkage analysis (co-inheritance of physically adjacent genes), described in, e.g., Egeland et al. (1987) *Nature* 325:783-787.

Moreover, differences in the DNA sequences between individuals affected and unaffected with a disease associated with a gene of the invention can be determined. If a mutation is observed in some or all of the affected individuals but not in any unaffected individuals, then the mutation is likely to be the causative agent of the particular disease. Comparison of affected and unaffected individuals generally involves first looking for structural alterations in the chromosomes such as deletions or translocations that are visible from chromosome spreads or detectable using PCR based on that DNA sequence. Ultimately, complete sequencing of genes from several individuals can be performed to confirm the presence of a mutation and to distinguish mutations from polymorphisms.

In the instant case, the human gene for TANGO 216 was mapped on radiation hybrid panels to the long arm of chromosome 4, in the region q11-13. Flanking markers for this region are GCT14E02 and jktbp-rs2. The JPD (periodontitis, juvenile), and DGI1(dentinogenesis imperfecta) loci also map to this region of the human chromosome. The GRO1 (FRO1 oncogene), ALB (albumin), IL8 (interleukin 8), HTN (histatin), and DCK (deoxycytidine kinase) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 5. The rs (recessive spotting) locus also maps to this region of the mouse chromosome. The ste (sulfotransferase), areg (amphiregulin), btc (betacellulin), mc (marcel), alb1 (albumin 1), and afp (alpha fetoprotein) genes also map to this region of the mouse chromosome.

In the instant case, the human gene for TANGO 261 was mapped on radiation hybrid panels to the long arm of chromosome 20, in the region q13.2-13.3 Flanking markers for this region are WI-3773 and AFMA202YB9. The EEGV1 (electroencephalographic variant pattern 1) and PHP1B (pseudohypoparathyroidism) loci also map to this region of the human chromosome. The MC3R (melanocortin 3 receptor), EDN3 (endothelin 3), ADA (adenosine deaminase), and OQTL (obesity QTL) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 2. The fc (flecking) and

ra (ragged) loci also map to this region of the mouse chromosome. The mc3r (melanocortin 3 receptor), fc (flecking), ra (ragged), and ntsr (neurotensin receptor) genes also map to this region of the mouse chromosome.

In the instant case, the human gene for TANGO 262 was mapped on radiation hybrid panels to the long arm of chromosome 14, in the region q23-q24. Flanking markers for this region are WI-6253 and WI-5815. The FNTB (fanesyltransferase) and MNAT1 (menage) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome 12.

In the instant case, the human gene for TANGO 267 was mapped on radiation hybrid panels to the long arm of human chromosome X, in the region q12. Flanking markers for this region are WI-5587 and WI5717. The Ar (androgen receptor), MSN (moesin), and OPHN (Oligophrenin I) genes also map to this region of the human chromosome. This region is syntenic to mouse chromosome. The gs (greasy) locus also maps to this region of the mouse chromosome. The ar (androgen receptor) and sla (sex linked anemia) genes also map to this region of the mouse chromosome.

A polypeptide and fragments and sequences thereof and antibodies specific thereto can be used to map the location of the gene encoding the polypeptide on a chromosome. This mapping can be carried out by specifically detecting the presence of the polypeptide in members of a panel of somatic cell hybrids between cells of a first species of animal from which the protein originates and cells from a second species of animal and then determining which somatic cell hybrid(s) expresses the polypeptide and noting the chromosome(s) from the first species of animal that it contains. For examples of this technique, see Pajunen *et al.* (1988) *Cytogenet. Cell Genet. 47*:37-41 and Van Keuren *et al.* (1986) *Hum. Genet. 74*:34-40. Alternatively, the presence of the polypeptide in the somatic cell hybrids can be determined by assaying an activity or property of the polypeptide, for example, enzymatic activity, as described in Bordelon-Riser *et al.* (1979) *Somatic Cell Genetics 5*:597-613 and Owerbach *et al.* (1978) *Proc. Natl. Acad. Sci. USA 75*:5640-5644.

2. Tissue Typing

The nucleic acid sequences of the present invention can also be used to identify individuals from minute biological samples. The United States military, for example, is considering the use of restriction fragment length polymorphism (RFLP) for identification of its personnel. In this technique, an individual's genomic DNA is digested with one or

more restriction enzymes, and probed on a Southern blot to yield unique bands for identification. This method does not suffer from the current limitations of "Dog Tags" which can be lost, switched, or stolen, making positive identification difficult. The sequences of the present invention are useful as additional DNA markers for RFLP (described in U.S. Patent 5,272,057).

Furthermore, the sequences of the present invention can be used to provide an alternative technique which determines the actual base-by-base DNA sequence of selected portions of an individual's genome. Thus, the nucleic acid sequences described herein can be used to prepare two PCR primers from the 5' and 3' ends of the sequences. These primers can then be used to amplify an individual's DNA and subsequently sequence it.

Panels of corresponding DNA sequences from individuals, prepared in this manner, can provide unique individual identifications, as each individual will have a unique set of such DNA sequences due to allelic differences. The sequences of the present invention can be used to obtain such identification sequences from individuals and from tissue. The nucleic acid sequences of the invention uniquely represent portions of the human genome. Allelic variation occurs to some degree in the coding regions of these sequences, and to a greater degree in the non-coding regions. It is estimated that allelic variation between individual humans occurs with a frequency of about once per each 500 bases. Each of the sequences described herein can, to some degree, be used as a standard against which DNA from an individual can be compared for identification purposes. Because greater numbers of polymorphisms occur in the noncoding regions, fewer sequences are necessary to differentiate individuals. The noncoding sequences of SEQ ID NO:1, 4, 7, 10, or 13 can comfortably provide positive individual identification with a panel of perhaps 10 to 1,000 primers which each yield a noncoding amplified sequence of 100 bases. If predicted coding sequences, such as those in SEQ ID NO:3, 6, 9, 12 or 15, are used, a more appropriate number of primers for positive individual identification would be 500-2,000.

If a panel of reagents from the nucleic acid sequences described herein is used to generate a unique identification database for an individual, those same reagents can later be used to identify tissue from that individual. Using the unique identification database, positive identification of the individual, living or dead, can be made from extremely small tissue samples.

3. Use of Partial Gene Sequences in Forensic Biology

DNA-based identification techniques can also be used in forensic biology. Forensic biology is a scientific field employing genetic typing of biological evidence found at a crime scene as a means for positively identifying, for example, a perpetrator of a crime. To make such an identification, PCR technology can be used to amplify DNA sequences taken from very small biological samples such as tissues, e.g., hair or skin, or body fluids, e.g., blood, saliva, or semen found at a crime scene. The amplified sequence can then be compared to a standard, thereby allowing identification of the origin of the biological sample.

The sequences of the present invention can be used to provide polynucleotide reagents, e.g., PCR primers, targeted to specific loci in the human genome, which can enhance the reliability of DNA-based forensic identifications by, for example, providing another "identification marker" (i.e. another DNA sequence that is unique to a particular individual). As mentioned above, actual base sequence information can be used for identification as an accurate alternative to patterns formed by restriction enzyme generated fragments. Sequences targeted to noncoding regions are particularly appropriate for this use as greater numbers of polymorphisms occur in the noncoding regions, making it easier to differentiate individuals using this technique. Examples of polynucleotide reagents include the nucleic acid sequences of the invention or portions thereof, e.g., fragments derived from noncoding regions having a length of at least 20 or 30 bases.

The nucleic acid sequences described herein can further be used to provide polynucleotide reagents, e.g., labeled or labelable probes which can be used in, for example, an *in situ* hybridization technique, to identify a specific tissue, e.g., brain tissue. This can be very useful in cases where a forensic pathologist is presented with a tissue of unknown origin. Panels of such probes can be used to identify tissue by species and/or by organ type.

C. Predictive Medicine

The present invention also pertains to the field of predictive medicine in which diagnostic assays, prognostic assays, pharmacogenomics, and monitoring clinical trails are used for prognostic (predictive) purposes to thereby treat an individual prophylactically. Accordingly, one aspect of the present invention relates to diagnostic assays for determining expression of a polypeptide or nucleic acid of the invention and/or activity of a polypeptide of the invention, in the context of a biological sample (e.g., blood, serum, cells, tissue) to thereby determine whether an individual is afflicted with a disease or disorder, or is at risk of developing a disorder, associated with aberrant expression or

activity of a polypeptide of the invention. The invention also provides for prognostic (or predictive) assays for determining whether an individual is at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, mutations in a gene of the invention can be assayed in a biological sample. Such assays can be used for prognostic or predictive purpose to thereby prophylactically treat an individual prior to the onset of a disorder characterized by or associated with aberrant expression or activity of a polypeptide of the invention.

Another aspect of the invention provides methods for expression of a nucleic acid or polypeptide of the invention or activity of a polypeptide of the invention in an individual to thereby select appropriate therapeutic or prophylactic agents for that individual (referred to herein as "pharmacogenomics"). Pharmacogenomics allows for the selection of agents (e.g., drugs) for therapeutic or prophylactic treatment of an individual based on the genotype of the individual (e.g., the genotype of the individual examined to determine the ability of the individual to respond to a particular agent).

Yet another aspect of the invention pertains to monitoring the influence of agents (e.g., drugs or other compounds) on the expression or activity of a polypeptide of the invention in clinical trials. These and other agents are described in further detail in the following sections.

1. Diagnostic Assays

An exemplary method for detecting the presence or absence of a polypeptide or nucleic acid of the invention in a biological sample involves obtaining a biological sample from a test subject and contacting the biological sample with a compound or an agent capable of detecting a polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention such that the presence of a polypeptide or nucleic acid of the invention is detected in the biological sample. A preferred agent for detecting mRNA or genomic DNA encoding a polypeptide of the invention is a labeled nucleic acid probe capable of hybridizing to mRNA or genomic DNA encoding a polypeptide of the invention. The nucleic acid probe can be, for example, a full-length cDNA, such as the nucleic acid of SEQ ID NO:1, 4, 7, 10, or 13, or a portion thereof, such as an oligonucleotide of at least 15, 30, 50, 100, 250 or 500 nucleotides in length and sufficient to specifically hybridize under stringent conditions to a mRNA or genomic DNA encoding a polypeptide of the invention. Other suitable probes for use in the diagnostic assays of the invention are described herein.

A preferred agent for detecting a polypeptide of the invention is an antibody capable of binding to a polypeptide of the invention, preferably an antibody with a detectable label. Antibodies can be polyclonal, or more preferably, monoclonal. An intact antibody, or a fragment thereof (e.g., Fab or $F(ab')_2$) can be used. The term "labeled", with regard to the probe or antibody, is intended to encompass direct labeling of the probe or antibody by coupling (i.e., physically linking) a detectable substance to the probe or antibody, as well as indirect labeling of the probe or antibody by reactivity with another reagent that is directly labeled. Examples of indirect labeling include detection of a primary antibody using a fluorescently labeled secondary antibody and end-labeling of a DNA probe with biotin such that it can be detected with fluorescently labeled streptavidin. The term "biological sample" is intended to include tissues, cells and biological fluids isolated from a subject, as well as tissues, cells and fluids present within a subject. That is, the detection method of the invention can be used to detect mRNA, protein, or genomic DNA in a biological sample *in vitro* as well as *in vivo*. For example, *in vitro* techniques for detection of mRNA include Northern hybridizations and *in situ* hybridizations. *In vitro* techniques for detection of a polypeptide of the invention include enzyme linked immunosorbent assays (ELISAs), Western blots, immunoprecipitations and immunofluorescence. *In vitro* techniques for detection of genomic DNA include Southern hybridizations. Furthermore, *in vivo* techniques for detection of a polypeptide of the invention include introducing into a subject a labeled antibody directed against the polypeptide. For example, the antibody can be labeled with a radioactive marker whose presence and location in a subject can be detected by standard imaging techniques.

In one embodiment, the biological sample contains protein molecules from the test subject. Alternatively, the biological sample can contain mRNA molecules from the test subject or genomic DNA molecules from the test subject. A preferred biological sample is a peripheral blood leukocyte sample isolated by conventional means from a subject.

In another embodiment, the methods further involve obtaining a control biological sample from a control subject, contacting the control sample with a compound or agent capable of detecting a polypeptide of the invention or mRNA or genomic DNA encoding a polypeptide of the invention, such that the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide is detected in the biological sample, and comparing the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the control sample with the presence of the polypeptide or mRNA or genomic DNA encoding the polypeptide in the test sample.

The invention also encompasses kits for detecting the presence of a polypeptide or nucleic acid of the invention in a biological sample (a test sample). Such kits can be used to determine if a subject is suffering from or is at increased risk of developing a disorder associated with aberrant expression of a polypeptide of the invention (e.g., a proliferative disorder, e.g., psoriasis or cancer). For example, the kit can comprise a labeled compound or agent capable of detecting the polypeptide or mRNA encoding the polypeptide in a biological sample and means for determining the amount of the polypeptide or mRNA in the sample (e.g., an antibody which binds the polypeptide or an oligonucleotide probe which binds to DNA or mRNA encoding the polypeptide). Kits may also include instructions for observing that the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide if the amount of the polypeptide or mRNA encoding the polypeptide is above or below a normal level.

For antibody-based kits, the kit may comprise, for example: (1) a first antibody (e.g., attached to a solid support) which binds to a polypeptide of the invention; and, optionally, (2) a second, different antibody which binds to either the polypeptide or the first antibody and is conjugated to a detectable agent.

For oligonucleotide-based kits, the kit may comprise, for example: (1) an oligonucleotide, e.g., a detectably labeled oligonucleotide, which hybridizes to a nucleic acid sequence encoding a polypeptide of the invention or (2) a pair of primers useful for amplifying a nucleic acid molecule encoding a polypeptide of the invention. The kit may also comprise, e.g., a buffering agent, a preservative, or a protein stabilizing agent. The kit may also comprise components necessary for detecting the detectable agent (e.g., an enzyme or a substrate). The kit may also contain a control sample or a series of control samples which can be assayed and compared to the test sample contained. Each component of the kit is usually enclosed within an individual container and all of the various containers are within a single package along with instructions for observing whether the tested subject is suffering from or is at risk of developing a disorder associated with aberrant expression of the polypeptide.

2. Prognostic Assays

The methods described herein can furthermore be utilized as diagnostic or prognostic assays to identify subjects having or at risk of developing a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example. the assays described herein, such as the preceding diagnostic assays or the

following assays, can be utilized to identify a subject having or at risk of developing a disorder associated with aberrant expression or activity of a polypeptide of the invention, e.g., a proliferative disorder, e.g., psoriasis or cancer, or an angiogenic disorder. Alternatively, the prognostic assays can be utilized to identify a subject having or at risk for developing such a disease or disorder. Thus, the present invention provides a method in which a test sample is obtained from a subject and a polypeptide or nucleic acid (e.g., mRNA, genomic DNA) of the invention is detected, wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject having or at risk of developing a disease or disorder associated with aberrant expression or activity of the polypeptide. As used herein, a "test sample" refers to a biological sample obtained from a subject of interest. For example, a test sample can be a biological fluid (e.g., serum), cell sample, or tissue.

Furthermore, the prognostic assays described herein can be used to determine whether a subject can be administered an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate) to treat a disease or disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, such methods can be used to determine whether a subject can be effectively treated with a specific agent or class of agents (e.g., agents of a type which decrease activity of the polypeptide). Thus, the present invention provides methods for determining whether a subject can be effectively treated with an agent for a disorder associated with aberrant expression or activity of a polypeptide of the invention in which a test sample is obtained and the polypeptide or nucleic acid encoding the polypeptide is detected (e.g., wherein the presence of the polypeptide or nucleic acid is diagnostic for a subject that can be administered the agent to treat a disorder associated with aberrant expression or activity of the polypeptide).

The methods of the invention can also be used to detect genetic lesions or mutations in a gene of the invention, thereby determining if a subject with the lesioned gene is at risk for a disorder characterized aberrant expression or activity of a polypeptide of the invention. In preferred embodiments, the methods include detecting, in a sample of cells from the subject, the presence or absence of a genetic lesion or mutation characterized by at least one of an alteration affecting the integrity of a gene encoding the polypeptide of the invention, or the mis-expression of the gene encoding the polypeptide of the invention. For example, such genetic lesions or mutations can be detected by ascertaining the existence of at least one of: 1) a deletion of one or more nucleotides from the gene; 2) an addition of one or more nucleotides to the gene; 3) a substitution of one or more nucleotides of the gene; 4) a chromosomal rearrangement of the gene; 5) an alteration in the level of a messenger RNA

transcript of the gene; 6) an aberrant modification of the gene, such as of the methylation pattern of the genomic DNA; 7) the presence of a non-wild type splicing pattern of a messenger RNA transcript of the gene; 8) a non-wild type level of a the protein encoded by the gene; 9) an allelic loss of the gene; and 10) an inappropriate post-translational modification of the protein encoded by the gene. As described herein, there are a large number of assay techniques known in the art which can be used for detecting lesions in a gene.

In certain embodiments, detection of the lesion involves the use of a probe/primer in a polymerase chain reaction (PCR) (*see, e.g.*, U.S. Patent Nos. 4,683,195 and 4,683,202), such as anchor PCR or RACE PCR, or, alternatively, in a ligation chain reaction (LCR) (*see, e.g.*, Landegran et al. (1988) *Science* 241:1077-1080; and Nakazawa et al. (1994) *Proc. Natl. Acad. Sci. USA* 91:360-364), the latter of which can be particularly useful for detecting point mutations in a gene (*see, e.g.*, Abravaya et al. (1995) *Nucleic Acids Res.* 23:675-682). This method can include the steps of collecting a sample of cells from a patient, isolating nucleic acid (e.g., genomic, mRNA or both) from the cells of the sample, contacting the nucleic acid sample with one or more primers which specifically hybridize to the selected gene under conditions such that hybridization and amplification of the gene (if present) occurs, and detecting the presence or absence of an amplification product, or detecting the size of the amplification product and comparing the length to a control sample. It is anticipated that PCR and/or LCR may be desirable to use as a preliminary amplification step in conjunction with any of the techniques used for detecting mutations described herein.

Alternative amplification methods include: self sustained sequence replication (Guatelli et al. (1990) *Proc. Natl. Acad. Sci. USA* 87:1874-1878), transcriptional amplification system (Kwoh, et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:1173-1177), Q-Beta Replicase (Lizardi et al. (1988) *Bio/Technology* 6:1197), or any other nucleic acid amplification method, followed by the detection of the amplified molecules using techniques well known to those of skill in the art. These detection schemes are especially useful for the detection of nucleic acid molecules if such molecules are present in very low numbers.

In an alternative embodiment, mutations in a selected gene from a sample cell can be identified by alterations in restriction enzyme cleavage patterns. For example, sample and control DNA is isolated, amplified (optionally), digested with one or more restriction endonucleases, and fragment length sizes are determined by gel electrophoresis

and compared. Differences in fragment length sizes between sample and control DNA indicates mutations in the sample DNA. Moreover, the use of sequence specific ribozymes (*see, e.g.*, U.S. Patent No. 5,498,531) can be used to score for the presence of specific mutations by development or loss of a ribozyme cleavage site.

In other embodiments, genetic mutations can be identified by hybridizing a sample and control nucleic acids, e.g., DNA or RNA, to high density arrays containing hundreds or thousands of oligonucleotides probes (Cronin et al. (1996) *Human Mutation* 7:244-255; Kozal et al. (1996) *Nature Medicine* 2:753-759). For example, genetic mutations can be identified in two-dimensional arrays containing light-generated DNA probes as described in Cronin et al., *supra*. Briefly, a first hybridization array of probes can be used to scan through long stretches of DNA in a sample and control to identify base changes between the sequences by making linear arrays of sequential overlapping probes. This step allows the identification of point mutations. This step is followed by a second hybridization array that allows the characterization of specific mutations by using smaller, specialized probe arrays complementary to all variants or mutations detected. Each mutation array is composed of parallel probe sets, one complementary to the wild-type gene and the other complementary to the mutant gene.

In yet another embodiment, any of a variety of sequencing reactions known in the art can be used to directly sequence the selected gene and detect mutations by comparing the sequence of the sample nucleic acids with the corresponding wild-type (control) sequence. Examples of sequencing reactions include those based on techniques developed by Maxim and Gilbert ((1977) *Proc. Natl. Acad. Sci. USA* 74:560) or Sanger ((1977) *Proc. Natl. Acad. Sci. USA* 74:5463). It is also contemplated that any of a variety of automated sequencing procedures can be utilized when performing the diagnostic assays ((1995) *Bio/Techniques* 19:448), including sequencing by mass spectrometry (*see, e.g.*, PCT Publication No. WO 94/16101; Cohen et al. (1996) *Adv. Chromatogr.* 36:127-162; and Griffin et al. (1993) *Appl. Biochem. Biotechnol.* 38:147-159).

Other methods for detecting mutations in a selected gene include methods in which protection from cleavage agents is used to detect mismatched bases in RNA/RNA or RNA/DNA heteroduplexes (Myers et al. (1985) *Science* 230:1242). In general, the technique of mismatch cleavage entails providing heteroduplexes formed by hybridizing (labeled) RNA or DNA containing the wild-type sequence with potentially mutant RNA or DNA obtained from a tissue sample. The double-stranded duplexes are treated with an agent which cleaves single-stranded regions of the duplex such as which will exist due to

basepair mismatches between the control and sample strands. RNA/DNA duplexes can be treated with RNase to digest mismatched regions, and DNA/DNA hybrids can be treated with S1 nuclease to digest mismatched regions.

In other embodiments, either DNA/DNA or RNA/DNA duplexes can be treated with hydroxylamine or osmium tetroxide and with piperidine in order to digest mismatched regions. After digestion of the mismatched regions, the resulting material is then separated by size on denaturing polyacrylamide gels to determine the site of mutation. *See, e.g.*, Cotton et al. (1988) *Proc. Natl. Acad. Sci. USA* 85:4397; Saleeba et al. (1992) *Methods Enzymol.* 217:286-295. In a preferred embodiment, the control DNA or RNA can be labeled for detection.

In still another embodiment, the mismatch cleavage reaction employs one or more proteins that recognize mismatched base pairs in double-stranded DNA (so called DNA mismatch repair enzymes) in defined systems for detecting and mapping point mutations in cDNAs obtained from samples of cells. For example, the mutY enzyme of E. coli cleaves A at G/A mismatches and the thymidine DNA glycosylase from HeLa cells cleaves T at G/T mismatches (Hsu et al. (1994) *Carcinogenesis* 15:1657-1662). According to an exemplary embodiment, a probe based on a selected sequence, e.g., a wild-type sequence, is hybridized to a cDNA or other DNA product from a test cell(s). The duplex is treated with a DNA mismatch repair enzyme, and the cleavage products, if any, can be detected from electrophoresis protocols or the like. *See, e.g.,* U.S. Patent No. 5,459,039.

In other embodiments, alterations in electrophoretic mobility will be used to identify mutations in genes. For example, single strand conformation polymorphism (SSCP) may be used to detect differences in electrophoretic mobility between mutant and wild type nucleic acids (Orita et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:2766; *see also* Cotton (1993) *Mutat. Res.* 285:125-144; Hayashi (1992) *Genet. Anal. Tech. Appl.* 9:73-79). Single-stranded DNA fragments of sample and control nucleic acids will be denatured and allowed to renature. The secondary structure of single-stranded nucleic acids varies according to sequence, and the resulting alteration in electrophoretic mobility enables the detection of even a single base change. The DNA fragments may be labeled or detected with labeled probes. The sensitivity of the assay may be enhanced by using RNA (rather than DNA), in which the secondary structure is more sensitive to a change in sequence. In a preferred embodiment, the subject method utilizes heteroduplex analysis to separate double stranded heteroduplex molecules on the basis of changes in electrophoretic mobility (Keen et al. (1991) *Trends Genet.* 7:5).

In yet another embodiment, the movement of mutant or wild-type fragments in polyacrylamide gels containing a gradient of denaturant is assayed using denaturing gradient gel electrophoresis (DGGE) (Myers et al. (1985) *Nature* 313:495). When DGGE is used as the method of analysis, DNA will be modified to insure that it does not completely denature, for example by adding a GC clamp of approximately 40 bp of high-melting GC-rich DNA by PCR. In a further embodiment, a temperature gradient is used in place of a denaturing gradient to identify differences in the mobility of control and sample DNA (Rosenbaum and Reissner (1987) *Biophys. Chem.* 265:12753).

Examples of other techniques for detecting point mutations include, but are not limited to, selective oligonucleotide hybridization, selective amplification, or selective primer extension. For example, oligonucleotide primers may be prepared in which the known mutation is placed centrally and then hybridized to target DNA under conditions which permit hybridization only if a perfect match is found (Saiki et al. (1986) *Nature* 324:163); Saiki et al. (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Such allele specific oligonucleotides are hybridized to PCR amplified target DNA or a number of different mutations when the oligonucleotides are attached to the hybridizing membrane and hybridized with labeled target DNA.

Alternatively, allele specific amplification technology which depends on selective PCR amplification may be used in conjunction with the instant invention. Oligonucleotides used as primers for specific amplification may carry the mutation of interest in the center of the molecule (so that amplification depends on differential hybridization) (Gibbs et al. (1989) *Nucleic Acids Res.* 17:2437-2448) or at the extreme 3' end of one primer where, under appropriate conditions, mismatch can prevent or reduce polymerase extension (Prossner (1993) *Tibtech* 11:238). In addition, it may be desirable to introduce a novel restriction site in the region of the mutation to create cleavage-based detection (Gasparini et al. (1992) *Mol. Cell Probes* 6:1). It is anticipated that in certain embodiments amplification may also be performed using Taq ligase for amplification (Barany (1991) *Proc. Natl. Acad. Sci. USA* 88:189). In such cases, ligation will occur only if there is a perfect match at the 3' end of the 5' sequence making it possible to detect the presence of a known mutation at a specific site by looking for the presence or absence of amplification.

The methods described herein may be performed, for example, by utilizing pre-packaged diagnostic kits comprising at least one probe nucleic acid or antibody reagent described herein, which may be conveniently used, e.g., in clinical settings to diagnose

patients exhibiting symptoms or family history of a disease or illness involving a gene encoding a polypeptide of the invention. Furthermore, any cell type or tissue, preferably peripheral blood leukocytes, in which the polypeptide of the invention is expressed may be utilized in the prognostic assays described herein.

3. Pharmacogenomics

Agents, or modulators which have a stimulatory or inhibitory effect on activity or expression of a polypeptide of the invention as identified by a screening assay described herein can be administered to individuals to treat (prophylactically or therapeutically) disorders associated with aberrant activity of the polypeptide. In conjunction with such treatment, the pharmacogenomics (i.e., the study of the relationship between an individual's genotype and that individual's response to a foreign compound or drug) of the individual may be considered. Differences in metabolism of therapeutics can lead to severe toxicity or therapeutic failure by altering the relation between dose and blood concentration of the pharmacologically active drug. Thus, the pharmacogenomics of the individual permits the selection of effective agents (e.g., drugs) for prophylactic or therapeutic treatments based on a consideration of the individual's genotype. Such pharmacogenomics can further be used to determine appropriate dosages and therapeutic regimens. Accordingly, the activity of a polypeptide of the invention, expression of a nucleic acid of the invention, or mutation content of a gene of the invention in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual.

Pharmacogenomics deals with clinically significant hereditary variations in the response to drugs due to altered drug disposition and abnormal action in affected persons. *See*, *e.g.*, Linder (1997) *Clin. Chem.* 43(2):254-266. In general, two types of pharmacogenetic conditions can be differentiated. Genetic conditions transmitted as a single factor altering the way drugs act on the body are referred to as "altered drug action." Genetic conditions transmitted as single factors altering the way the body acts on drugs are referred to as "altered drug metabolism". These pharmacogenetic conditions can occur either as rare defects or as polymorphisms. For example, glucose-6-phosphate dehydrogenase deficiency (G6PD) is a common inherited enzymopathy in which the main clinical complication is haemolysis after ingestion of oxidant drugs (anti-malarials, sulfonamides, analgesics, nitrofurans) and consumption of fava beans.

As an illustrative embodiment, the activity of drug metabolizing enzymes is a major determinant of both the intensity and duration of drug action. The discovery of genetic polymorphisms of drug metabolizing enzymes (e.g., N-acetyltransferase 2 (NAT 2) and cytochrome P450 enzymes CYP2D6 and CYP2C19) has provided an explanation as to why some patients do not obtain the expected drug effects or show exaggerated drug response and serious toxicity after taking the standard and safe dose of a drug. These polymorphisms are expressed in two phenotypes in the population, the extensive metabolizer (EM) and poor metabolizer (PM). The prevalence of PM is different among different populations. For example, the gene coding for CYP2D6 is highly polymorphic and several mutations have been identified in PM, which all lead to the absence of functional CYP2D6. Poor metabolizers of CYP2D6 and CYP2C19 quite frequently experience exaggerated drug response and side effects when they receive standard doses. If a metabolite is the active therapeutic moiety, a PM will show no therapeutic response, as demonstrated for the analgesic effect of codeine mediated by its CYP2D6-formed metabolite morphine. The other extreme are the so called ultra-rapid metabolizers who do not respond to standard doses. Recently, the molecular basis of ultra-rapid metabolism has been identified to be due to CYP2D6 gene amplification

Thus, the activity of a polypeptide of the invention, expression of a nucleic acid encoding the polypeptide, or mutation content of a gene encoding the polypeptide in an individual can be determined to thereby select appropriate agent(s) for therapeutic or prophylactic treatment of the individual. In addition, pharmacogenetic studies can be used to apply genotyping of polymorphic alleles encoding drug-metabolizing enzymes to the identification of an individual's drug responsiveness phenotype. This knowledge, when applied to dosing or drug selection, can avoid adverse reactions or therapeutic failure and thus enhance therapeutic or prophylactic efficiency when treating a subject with a modulator of activity or expression of the polypeptide, such as a modulator identified by one of the exemplary screening assays described herein.

4. <u>Monitoring of Effects During Clinical Trials</u>

Monitoring the influence of agents (e.g., drugs, compounds) on the expression or activity of a polypeptide of the invention (e.g., the ability to modulate aberrant cell proliferation chemotaxis, and/or differentiation) can be applied not only in basic drug screening, but also in clinical trials. For example, the effectiveness of an agent, as determined by a screening assay as described herein, to increase gene expression, protein

levels or protein activity, can be monitored in clinical trials of subjects exhibiting decreased gene expression, protein levels, or protein activity. Alternatively, the effectiveness of an agent, as determined by a screening assay, to decrease gene expression, protein levels or protein activity, can be monitored in clinical trials of subjects exhibiting increased gene expression, protein levels, or protein activity. In such clinical trials, expression or activity of a polypeptide of the invention and preferably, that of other polypeptide that have been implicated in for example, a cellular proliferation disorder, can be used as a marker of the immune responsiveness of a particular cell.

For example, and not by way of limitation, genes, including those of the invention, that are modulated in cells by treatment with an agent (e.g., compound, drug or small molecule) which modulates activity or expression of a polypeptide of the invention (e.g., as identified in a screening assay described herein) can be identified. Thus, to study the effect of agents on cellular proliferation disorders, for example, in a clinical trial, cells can be isolated and RNA prepared and analyzed for the levels of expression of a gene of the invention and other genes implicated in the disorder. The levels of gene expression (i.e., a gene expression pattern) can be quantified by Northern blot analysis or RT-PCR, as described herein, or alternatively by measuring the amount of protein produced, by one of the methods as described herein, or by measuring the levels of activity of a gene of the invention or other genes. In this way, the gene expression pattern can serve as a marker, indicative of the physiological response of the cells to the agent. Accordingly, this response state may be determined before, and at various points during, treatment of the individual with the agent.

In a preferred embodiment, the present invention provides a method for monitoring the effectiveness of treatment of a subject with an agent (e.g., an agonist, antagonist, peptidomimetic, protein, peptide, nucleic acid, small molecule, or other drug candidate identified by the screening assays described herein) comprising the steps of (i) obtaining a pre-administration sample from a subject prior to administration of the agent; (ii) detecting the level of the polypeptide or nucleic acid of the invention in the preadministration sample; (iii) obtaining one or more post-administration samples from the subject; (iv) detecting the level the of the polypeptide or nucleic acid of the invention in the post-administration samples; (v) comparing the level of the polypeptide or nucleic acid of the invention in the pre-administration sample with the level of the polypeptide or nucleic acid of the invention in the post-administration sample or samples; and (vi) altering the administration of the agent to the subject accordingly. For example, increased administration of the agent may be desirable to increase the expression or activity of the

polypeptide to higher levels than detected, i.e., to increase the effectiveness of the agent. Alternatively, decreased administration of the agent may be desirable to decrease expression or activity of the polypeptide to lower levels than detected, i.e., to decrease the effectiveness of the agent.

C. Methods of Treatment

The present invention provides for both prophylactic and therapeutic methods of treating a subject at risk of (or susceptible to) a disorder or having a disorder associated with aberrant expression or activity of a polypeptide of the invention. For example, disorders characterized by abberant expression or activity of the polypeptides of the invention include proliferative disorders such as cancer.

1. Prophylactic Methods

In one aspect, the invention provides a method for preventing in a subject, a disease or condition associated with an aberrant expression or activity of a polypeptide of the invention, by administering to the subject an agent which modulates expression or at least one activity of the polypeptide. Subjects at risk for a disease which is caused or contributed to by aberrant expression or activity of a polypeptide of the invention can be identified by, for example, any or a combination of diagnostic or prognostic assays as described herein. Administration of a prophylactic agent can occur prior to the manifestation of symptoms characteristic of the aberrancy, such that a disease or disorder is prevented or, alternatively, delayed in its progression. Depending on the type of aberrancy, for example, an agonist or antagonist agent can be used for treating the subject. The appropriate agent can be determined based on screening assays described herein.

2. Therapeutic Methods

Another aspect of the invention pertains to methods of modulating expression or activity of a polypeptide of the invention for therapeutic purposes. The modulatory method of the invention involves contacting a cell with an agent that modulates one or more of the activities of the polypeptide. An agent that modulates activity can be an agent as described herein, such as a nucleic acid or a protein, a naturally-occurring cognate ligand of the polypeptide, a peptide, a peptidomimetic, or other small molecule. In one embodiment, the agent stimulates one or more of the biological activities of the polypeptide.

Examples of such stimulatory agents include the active polypeptide of the invention and a nucleic acid molecule encoding the polypeptide of the invention that has been introduced into the cell. In another embodiment, the agent inhibits one or more of the biological activities of the polypeptide of the invention. Examples of such inhibitory agents include antisense nucleic acid molecules and antibodies. These modulatory methods can be performed *in vitro* (e.g., by culturing the cell with the agent) or, alternatively, *in vivo* (e.g, by administering the agent to a subject). As such, the present invention provides methods of treating an individual afflicted with a disease or disorder characterized by aberrant expression or activity of a polypeptide of the invention. In one embodiment, the method involves administering an agent (e.g., an agent identified by a screening assay described herein), or combination of agents that modulates (e.g., upregulates or downregulates) expression or activity. In another embodiment, the method involves administering a polypeptide of the invention or a nucleic acid molecule of the invention as therapy to compensate for reduced or aberrant expression or activity of the polypeptide.

Stimulation of activity is desirable in situations in which activity or expression is abnormally low or downregulated and/or in which increased activity is likely to have a beneficial effect. Conversely, inhibition of activity is desirable in situations in which activity or expression is abnormally high or upregulated and/or in which decreased activity is likely to have a beneficial effect.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

Deposit of Clones

Clones containing cDNA molecules encoding human TANGO 216, TANGO 261, TANGO 262, TANGO 266, and TANGO 267 (clones EpT216, EpT261, EpT262, EpT266, and EpT267, respectively), were deposited with the American Type Culture Collection (Manassas, VA) on March 26, 1999 as Accession Number 207176, as part of a composite deposit representing a mixture of five strains, each carrying one recombinant plasmid harboring a particular cDNA clone.

To distinguish the strains and isolate a strain harboring a particular cDNA clone, an aliquot of the mixture can be streaked out to single colonies on nutrient medium (e.g., LB plates) supplemented with 100μg/ml ampicillin, single colonies grown, and then

plasmid DNA extracted using a standard minipreparation procedure. Next, a sample of the DNA minipreparation can be digested with a combination of the restriction enzymes *Sal* I and *Not* I and the resultant products resolved on a 0.8% agarose gel using standard DNA electrophoresis conditions. The digest liberates fragments as follows:

TANGO 216 (EpT216): 4.4 kb

TANGO 261 (EpT261): 1.9 kb

TANGO 262 (EpT262): 1.5 kb

TANGO 266 (EpT266): .4 kb

TANGO 267 (EpT267): 2.8 kb

The identity of the strains can be inferred from the fragments liberated.

All publications, patents and patent applications mentioned in this specification are herein incorporated by reference in to the specification to the same extent as if each individual publication, patent or patent application was specifically and individually indicated to be incorporated herein by reference.

Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

International Application No: PCT/

MICROORGANISMS

Optional Sheet in connection with the microorganism referred to on pages 129, lines 25-32 of the description '

A. IDENTIFICATION OF DEPOSIT '

Further deposits are identified on an additional sheet '

Name of depositary institution '

American Type Culture Collection

Address of depositary institution (including postal code and country) '

10801 University Blvd.
Manassas, VA 20110-2209
US

Date of deposit ' March 26, 1999 Accession Number ' 207176

B. ADDITIONAL INDICATIONS ' (leave blank if not applicable). This information is continued on a separate attached sheet

C. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE ' (if the indications are not all designated States)

D. SEPARATE FURNISHING OF INDICATIONS ' (leave blank if not applicable)

The indications listed below will be submitted to the International Bureau later ' (Specify the general nature of the indications e.g., "Accession Number of Deposit")

E. ☑ This sheet was received with the International application when filed (to be checked by the receiving Office)

Marilyn Younger
(Authorized Officer)

☐ The date of receipt (from the applicant) by the International Bureau '

was

(Authorized Officer)

What is claimed is:

1. An isolated nucleic acid molecule selected from the group consisting of:

a) a nucleic acid molecule comprising a nucleotide sequence which is at least 55% identical to the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, or the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, or a complement thereof;

b) a nucleic acid molecule comprising a fragment of at least 300 nucleotides of the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24 or the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, or a complement thereof;

c) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176;

d) a nucleic acid molecule which encodes a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the polypeptide encoded by the cDNA insert of the plasmid deposited with ATCC™ as Accession Number 207176, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or thepolypeptide encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176; and

e) a nucleic acid molecule which encodes a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, wherein the nucleic acid molecule hybridizes to a nucleic acid molecule comprising SEQ ID NO:3, 6, 9, 12, 15, 18, 21, 24 or a complement thereof under stringent conditions.

2. The isolated nucleic acid molecule of claim 1, which is selected from the group consisting of:

a) a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 3, 4, 6, 7, 9, 10, 12, 13, 15, 16, 18, 19, 21, 22, 24, or the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, or a complement thereof; and

b) a nucleic acid molecule which encodes a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176.

3. The nucleic acid molecule of claim 1 further comprising vector nucleic acid sequences.

4. The nucleic acid molecule of claim 1 further comprising nucleic acid sequences encoding a heterologous polypeptide.

5. A host cell which contains the nucleic acid molecule of claim 1.

6. The host cell of claim 5 which is a mammalian host cell.

7. A non-human mammalian host cell containing the nucleic acid molecule of claim 1.

8. An isolated polypeptide selected from the group consisting of:

a) a fragment of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23.

b) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 4, 7, 10, 13, 16, 19, 22 or a complement thereof under stringent conditions; and

c) a polypeptide which is encoded by a nucleic acid molecule comprising a nucleotide sequence which is at least 65% identical to a nucleic acid comprising the nucleotide sequence of SEQ ID NO:1, 4, 7, 8, 10, 13, or a complement thereof.

9. The isolated polypeptide of claim 8 comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, or 23.

10. The polypeptide of claim 8 further comprising heterologous amino acid sequences.

11. An antibody which selectively binds to a polypeptide of claim 8.

12. A method for producing a polypeptide selected from the group consisting of:

a) a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176;

b) a polypeptide comprising a fragment of the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, wherein the fragment comprises at least 15 contiguous amino acids of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176; and

c) a naturally occurring allelic variant of a polypeptide comprising the amino acid sequence of SEQ ID NO:2, 5, 8, 11, 14, 17, 20, 23, or the amino acid sequence encoded by the cDNA insert of the plasmid deposited with the ATCC™ as Accession Number 207176, wherein the polypeptide is encoded by a nucleic acid molecule which hybridizes to a nucleic acid molecule comprising SEQ ID NO:1, 4, 7, 10, 13, 16, 19, 22, or a complement thereof under stringent conditions;

comprising culturing the host cell of claim 5 under conditions in which the nucleic acid molecule is expressed.

13. A method for detecting the presence of a polypeptide of claim 8 in a sample, comprising:

a) contacting the sample with a compound which selectively binds to a polypeptide of claim 8; and

b) determining whether the compound binds to the polypeptide in the sample.

14. The method of claim 13, wherein the compound which binds to the polypeptide is an antibody.

15. A kit comprising a compound which selectively binds to a polypeptide of claim 8 and instructions for use.

16. A method for detecting the presence of a nucleic acid molecule of claim 1 in a sample, comprising the steps of:

a) contacting the sample with a nucleic acid probe or primer which selectively hybridizes to the nucleic acid molecule; and

b) determining whether the nucleic acid probe or primer binds to a nucleic acid molecule in the sample.

17. The method of claim 16, wherein the sample comprises mRNA molecules and is contacted with a nucleic acid probe.

18. A kit comprising a compound which selectively hybridizes to a nucleic acid molecule of claim 1 and instructions for use.

19. A method for identifying a compound which binds to a polypeptide of claim 8 comprising the steps of:

a) contacting a polypeptide, or a cell expressing a polypeptide of claim 8 with a test compound; and

b) determining whether the polypeptide binds to the test compound.

20. The method of claim 19, wherein the binding of the test compound to the polypeptide is detected by a method selected from the group consisting of:

a) detection of binding by direct detecting of test compound/polypeptide binding;

b) detection of binding using a competition binding assay;

c) detection of binding using an assay for TANGO 216, TANGO 261, TANGO 262, TANGO 266 or TANGO 267-mediated signal transduction.

21. A method for modulating the activity of a polypeptide of claim 8 comprising contacting a polypeptide or a cell expressing a polypeptide of claim 8 with a compound which binds to the polypeptide in a sufficient concentration to modulate the activity of the polypeptide.

22. A method for identifying a compound which modulates the activity of a polypeptide of claim 8, comprising:

a) contacting a polypeptide of claim 8 with a test compound; and

b) determining the effect of the test compound on the activity of the polypeptide to thereby identify a compound which modulates the activity of the polypeptide.

CTCCACTCCTAGTTTGTTCTGCCGTCGCCGCGTCCCAGGGACCCCTTGTCCCGAAGCGCACG 62

GCAGCGGGGGGGACTTCAGCCCTCCAGGCGGGGTGGGTTCCAGGTCCGGGTCCGAGGCGGGCGCTGGAGGCTCGGCCCC 141

AGGCCGGAGAGGAACTCCTTTCGCGAGCTGCCGCAGGCGCCGGCGTCTCAGCTGCTCGCCGCCCCCCACCCCAGAGTGC 220

GTGCCGGGTGACTCCCGCCACCTTTGCGACCCTCCTGAGCTTAGGGGACTGCGAGCGGGAGGGAGTCTCAGGCCCCCGG 299

M V A E R S P A R S P G S W L F P G 18
CCGCAGG ATG GTG GCG GAG CGG TCC CCG GCC CGC AGC CCC GGG AGC TGG CTG TTC CCC GGG 360

L W L L V L S G P G G L L R A Q E Q P S 38
CTG TGG CTG TTG GTG CTC AGC GGT CCC GGG GGG CTG CTG CGC GCC CAG GAG CAG CCC TCC 420

C R R A F D L Y F V L D K S G S V A N N 58
TGC AGA AGA GCC TTT GAT CTC TAC TTC GTC CTG GAC AAG TCT GGG AGT GTG GCA AAT AAC 480

W I E I Y N F V Q Q L A E R F V S P E M 78
TGG ATT GAA ATT TAT AAT TTC GTA CAG CAA CTT GCG GAG AGA TTT GTG AGC CCT GAA ATG 540

R L S F I V F S S Q A T I I L P L T G D 98
AGA TTA TCT TTC ATT GTG TTT TCT TCT CAA GCA ACT ATT ATT TTG CCA TTA ACT GGA GAC 600

R G K I S K G L E D L K R V S P V G E T 118
AGA GGC AAA ATC AGT AAA GGC TTG GAG GAT TTA AAA CGT GTT AGT CCA GTA GGA GAG ACA 660

Y I H E G L K L A N E Q I Q K A G G L K 138
TAT ATC CAT GAA GGA CTA AAG CTA GCG AAT GAA CAA ATT CAG AAA GCA GGA GGC TTG AAA 720

T S S I I I A L T D G K L D G L V P S Y 158
ACC TCC AGT ATC ATA ATT GCT CTG ACA GAT GGC AAG TTG GAC GGT CTG GTG CCA TCA TAT 780

A E K E A K I S R S L G A S V Y C V G V 178
GCA GAG AAA GAG GCA AAG ATA TCC AGG TCA CTT GGG GCT AGT GTT TAT TGT GTT GGT GTC 840

L D F E Q A Q L E R I A D S K E Q V F P 198
CTT GAT TTT GAA CAA GCA CAG CTT GAA AGA ATT GCT GAT TCC AAG GAG CAA GTT TTC CCT 900

V K G G F Q A L K G I I N S I L A Q S C 218
GTC AAA GGT GGA TTT CAG GCT CTT AAA GGA ATA ATT AAT TCT ATA CTA GCT CAG TCA TGT 960

T E I L E L Q P S S V C V G E E F Q I V 238
ACT GAA ATC CTA GAA TTG CAG CCC TCA AGT GTC TGT GTG GGG GAG GAA TTT CAG ATT GTC 1020

L S G R G F M L G S R N G S V L C T Y T 258
TTA AGT GGA AGA GGA TTC ATG CTG GGC AGT CGG AAT GGC AGT GTT CTC TGC ACT TAC ACT 1080

FIG.1A

V N E T Y T T S V K P V S V Q L N S M L 278
GTA AAT GAA ACA TAT ACA ACG AGT GTA AAA CCA GTA AGT GTA CAG CTT AAT TCT ATG CTT 1140

C P A P I L N K A G E T L D V S V S F N 298
TGT CCT GCA CCT ATC CTG AAT AAA GCT GGA GAA ACT CTT GAT GTT TCA GTG AGC TTT AAT 1200

G G K S V I S G S L I V T A T E C S N G 318
GGA GGA AAA TCT GTC ATT TCA GGA TCA TTA ATT GTC ACA GCC ACA GAA TGT TCT AAC GGG 1260

I A A I I V I L V L L L L L G I G L M W 338
ATC GCA GCC ATC ATT GTT ATT TTG GTG TTA CTG CTA CTC CTG GGG ATC GGT TTG ATG TGG 1320

W F W P L C C K V V I K D P P P P P P P 358
TGG TTT TGG CCC CTT TGC TGC AAA GTG GTT ATT AAG GAT CCT CCA CCA CCA CCC CCC CCT 1380

A P K E E E E E P L P T K K W P T V D A 378
GCA CCA AAA GAG GAG GAA GAA GAA CCT TTG CCT ACT AAA AAG TGG CCA ACT GTG GAT GCT 1440

S Y Y G G R G V G G I K R M E V R W G D 398
TCC TAT TAT GGT GGT CGA GGG GTT GGA GGA ATT AAA AGA ATG GAG GTT CGT TGG GGT GAT 1500

K G S T E E G A R L E K A K N A V V K I 418
AAA GGA TCT ACT GAG GAA GGT GCA AGG CTA GAG AAA GCC AAA AAT GCT GTG GTG AAG ATT 1560

P E E T E E P I R P R P P R P K P T H Q 438
CCT GAA GAA ACA GAG GAA CCC ATC AGG CCT AGA CCA CCT CGA CCC AAA CCC ACA CAC CAG 1620

P P Q T K W Y T P I K G R L D A L W A L 458
CCT CCT CAG ACA AAA TGG TAC ACC CCA ATT AAG GGT CGT CTT GAT GCT CTC TGG GCT TTG 1680

L R R Q Y D R V S L M R P Q E G D E G R 478
TTG AGG CGG CAG TAT GAC CGG GTT TCT TTG ATG CGA CCT CAG GAA GGA GAT GAG GGC CGG 1740

C I N F S R V P S Q * 489
TGC ATA AAC TTC TCC CGA GTT CCA TCT CAG TAA 1773

AAGGGAAGCAGGAAGACCAAGAAGGTACGAAGATGGCACATTTTCACATAGCTGATTTTCAACCAAATGAAAAAAATCA 1852

AGTGCATTTCAGAAGCTTTTGGAAGAGCAGCTTAATTCCTCTCAGTCGGGAAATGTTTTCTCTGCCTTCTGCTTTGCTT 1931

GCACCAAACATTTCTAAACACTTGTTCTGCCATCTACATGGGAGGTGATGAAACTCAGTGGTAACTCATGATTTATGAC 2010

ATTGAAAATAAAGAGGAACATTGACCTGCAGACTATGGTTTGTACAAGAAAGTTTGTTTGAATGTGTAGAACGAGGAAA 2089

AAGCAACAACAGCAACAACACGAAGATGATACCAAAACAAGGACCACAAAACAACTAGCCATGATGGGAGACAGGAGTT 2168

FIG.1B

TTTTACATGGAAACATGGCACTTGTGTTTTTATGTGGCAAGATCTTTATCCATAGGCAGAGTATGAAATTTCCCACCAG 2247

GCTAAGCAAATAAAGAAGTCCATTGCCTTATAGCTATGTCAGATCACAGAATCCTTCCAAGTGCTCTATCACAGTGTGC 2326

CTTATGGGAAGTTTCTGACTGGAAAATCTTGTCATTCTAACACTGAAAAGTGCACACGCATGACAAAATGTAGACAAGA 2405

TGCCTCAAGGTATTGGTAGCAAGCAAGATTTTGCCCTTTAGTTTTCGAAGACACCTTTCTTTCATTATGCACTCGGGAC 2484

AAGAAAATTAATAGAGCGTTATTCCACAGAAGGCCTCTAGCCAGAGATCTTGAGTGTAGTGCAAGGGACTCATGCTTTG 2563

CGAACTTGTCCCTGTGACTAGTAGATTCCCCCTTTTCCTGTGTTTAGGATTTAGTAGTGCATAAAGCATTAATATCCAT 2642

AAACATACCTAGAAGTTTGTTTTGCTTTTAATTTAAAGGAAGCAGTAACCACAAAGCTTCCGCTCAGGGTTTTTTCTTT 2721

CTTCAAGTCTCCAAGGGCTCTTCAGCGTCACAAGCCAGCAACTCTCTTTGCATTAAAATTTCAAAGTTTAATTAATATA 2800

ATTAAAAGCAACAGCAAGCAGCAGCCTGTGAAGATTTTGCTCATCTTTTTTATGCCTTTTGACATTGAATGACCTATTA 2879

CTGTATGCGCATTACTTGGATTTTGAGGGGCACTCTACCTTGGTTATGATTCAGTAGAGGAAAAAGACCACCTTTCTTC 2958

AATTTACAAATTAAATCTTCTGGAGGGTCGCTATCACAAAACATTGACGATGTATGTATTATAATTTTTTAGAAAAACC 3038

ACCATCGTGTCACGTCGACGATGCCAAATTATGTTAGCGTGAGCAGAAACACCGTGGGGGAGGAAGGCAGCAGCTGAAG 3116

AAAAAAGCTCAAATGATCTAGTCACTTTCGATACTGTACTTCAGATGCGAAATGGATATTCGAGTGGAAACCTGACAAA 3195

GTGCGCCTGCTTTGATGTGAACTGGTATAGACAATGACCAGTGGCTGGGTCAGTGGGATGTCTCTCTGTGAGCACAAAG 3274

GCTTATCAAATGACACTAAAGATAAGTTCAACAACCATCACATTGGAAGGGAGAAGGCGAACATTTCATGTTTGGCGGG 3353

CATGTGAGTGCACAAGATGGAAAGAGCGATTGGAGCATCCTGGTATAATTACCCCCATTGTGCTCTTAATGGAAATTTC 3432

AAAGGACGGGAGTATTCTGTTGGTTGGTGTCCAGGTTTGTGGCACTGTTCCAAGAGGCCTTACACACACACACAAATAT 3511

ATAATTTTCTATACATATATATCCTCTAGCTTGAAACTTTTGCTCAAGTTTATTTATGTCCCTGGCTGGCTGGATCCAA 3590

AGTCATGTGTCCACACATTCATAAATAAAAATTTTACCTATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3669

AAAAAAAA 3677

FIG.1C

TTCTCGCGAGCTGCCGCAGGCGCCGGCGTCTCAGCTGCTCGCCCGCCCCCCACCCCAGAGTGCGTGCCGGGTGACTCCCG 79

M V 2
CCACCTTTGCGACCCTCCTGAGCTTAGGGGACTGCGAGCGGGAGCGCGTCTCAGGCCCCCGGCCCGCAGG ATG GTG 154

A G R S R A R S P G S W L F P G L W L L 22
GCC GGT CGG TCC CGG GCG CGC AGC CCT GGG AGC TGG CTG TTC CCT GGC CTG TGG TTG TTG 214

A V G G P G S L L Q A Q E Q P S C K K A 42
GCT GTG GGC GGT CCG GGG TCG TTG CTG CAA GCC CAG GAG CAG CCC TCT TGC AAA AAA GCC 274

F D L Y F V L D K S G S V A N N W I E I 62
TTC GAT TTG TAC TTC GTA CTG GAC AAG TCT GGC AGT GTA GCA AAT AAC TGG ATT GAA ATT 334

Y N F V H Q L T E R F V S P E M R L S F 82
TAT AAT TTT GTC CAC CAG CTG ACA GAG AGA TTT GTG AGC CCT GAA ATG AGA TTG TCC TTC 394

I V F S S Q A T I I L P L T G D R Y K I 102
ATT GTG TTT TCT TCC CAA GCA ACC ATT ATT TTG CCA TTA ACT GGA GAC AGG TAC AAA ATT 454

G K G L E D L K A V K P V G E T Y I H E 122
GGC AAA GGA CTG GAG GAT TTA AAG GCC GTT AAG CCA GTT GGA GAA ACA TAC ATC CAT GAA 514

G L K L A N E Q I Q N A G G L K A S S I 142
GGA CTA AAG CTT GCA AAC GAA CAA ATT CAA AAT GCA GGA GGC TTA AAA GCC TCC AGT ATC 574

I I A L T D G K L D G L V P S Y A E N E 162
ATA ATT GCT TTG ACG GAC GGT AAG CTG GAC GGC CTG GTA CCA TCT TAT GCA GAG AAC GAG 634

A K K S R S L G A S V Y C V G V L D F E 182
GCA AAG AAG TCC AGG TCA CTT GGC GCT AGT GTT TAC TGC GTT GGG GTC CTT GAT TTT GAA 694

Q A Q L E R I A D S K D Q V F P V K G G 202
CAA GCT CAG CTG GAA AGA ATT GCT GAT TCC AAG GAC CAG GTT TTC CCT GTC AAA GGT GGA 754

F Q A L K G I I N S I L A Q S C T E I L 222
TTT CAA GCT CTC AAA GGC ATC ATC AAC TCT ATA TTA GCT CAA TCA TGT ACT GAA ATC CTG 814

E L S P S S V C V G E K F Q V V L T G R 242
GAA TTG AGT CCT TCA AGT GTC TGT GTA GGG GAG AAA TTT CAA GTT GTT CTG ACT GGA AGA 874

A V T S I S H D G S V L C T F T A N S T 262
GCA GTC ACG TCG ATC AGT CAC GAT GGC AGT GTC CTC TGT ACA TTC ACT GCA AAC AGC ACA 934

Y T K S E K P V S I Q P S S I L C P A P 282
TAT ACA AAG AGT GAG AAG CCA GTG AGC ATT CAG CCA AGT TCC ATC CTT TGT CCT GCA CCT 994

FIG.2A

```
V   L   N   K   D   G   E   T   L   E   V   S   I   S   Y   N   D   G   K   S    302
GTC CTG AAC AAA GAT GGA GAA ACT CTT GAA GTT TCA ATC AGC TAT AAT GAT GGG AAG TCT 1054

A   V   S   R   S   L   T   I   T   A   T   E   C   T   N   G   I   A   A   I    322
GCT GTC TCA AGA TCC TTA ACA ATC ACA GCC ACA GAA TGT ACC AAT GGG ATT GCA GCC ATC 1114

V   A   I   L   V   L   L   L   L   L   G   A   A   L   M   W   W   F   W   P    342
GTA GCT ATT TTG GTG TTG TTG CTG CTC TTG GGT GCT GCC TTG ATG TGG TGG TTT TGG CCC 1174

L   C   C   K   V   V   I   K   D   P   P   P   P   P   S   A   P   M   E   E    362
CTT TGC TGC AAA GTG GTT ATC AAG GAC CCT CCC CCA CCA CCT TCT GCA CCA ATG GAG GAG 1234

E   E   E   D   P   L   P   N   K   K   W   P   T   V   D   A   S   Y   Y   G    382
GAG GAG GAG GAT CCT TTG CCC AAC AAG AAG TGG CCG ACT GTG GAC GCT TCC TAC TAC GGA 1294

G   R   G   V   G   G   I   K   R   M   E   V   R   W   G   D   K   G   S   T    402
GGT CGA GGT GTT GGA GGA ATT AAA AGG ATG GAG GTC CGC TGG GGA GAT AAA GGA TCT ACA 1354

E   E   G   A   R   L   E   K   A   K   N   A   V   V   M   V   P   E   E   E    422
GAG GAA GGT GCA AGG CTA GAG AAA GCA AAA AAT GCT GTG GTG ATG GTC CCT GAA GAA GAG 1414

I   P   I   P   S   R   P   P   R   P   R   P   T   H   Q   A   P   Q   T   K    442
ATC CCC ATC CCA TCC AGA CCA CCT CGA CCC AGA CCC ACA CAC CAG GCA CCT CAG ACA AAG 1474

W   Y   T   P   I   K   G   R   L   D   A   L   W   A   L   I   M   K   Q   Y    462
TGG TAC ACG CCA ATC AAG GGT CGT TTG GAT GCA CTC TGG GCT TTG ATC ATG AAG CAG TAT 1534

D   R   V   S   L   M   R   P   Q   E   G   D   E   G   R   C   I   N   F   S    482
GAC CGG GTG TCC TTG ATG AGA CCC CAG GAA GGT GAT GAG GGC CGA TGC ATA AAC TTC TCC 1594

R   V   P   H   Q   *                                                            488
CGG GTT CCA CAT CAA TAA                                                         1612

GACGGGAGAACAAAGAATGAGAAGATAAGAAGACAGTGTGACAGTGTATCTTCATGATGCTGATTTCCAACAGAACCGA 1691

CAGTCCGGTGCATCTCAGAAGTTCTTGGGACAACAGCCCATTTTCTCTCTGTCAGGAAATGTTTCCTCTGCCTTCTGCT 1770

TTCTGTGCACTAAACATTTTCCAACACTTGTTCTGCCATCGACATGAGAGGTGATGAAAGTCATCGGCGATAGCCCATG 1849

ATTCACGACACTGAAAATCCCGAGGAATGTTAGTTTGCATGCTAGGGTTTATGCAAAGCTCGTTTTGACTATGTAAGAG 1928

GACAAAGCAGACACATCGATGATGTAATGATACCAAAGCTAGGACTGCAAATCCACCAGCCACAGAGGTTTGCAATGGA 2007

GACTGGTGATTCTGCCATGAATGTGTGGCCCCTGTGCTTTTGTTTGGCAAGATCTTTAGCTACAAGCAAAACATGAAGT 2086

TTCCTCCCAGGCTAAACAGATAATGGAGTCCACTGCCTTGTAGCTATGTCAGATAGCAAAGCCTTTCCAAGTCCTCCAT 2165
```

FIG.2B

TACTTTGTGCCTTACAGGAAATTTCTGACTAGAAAATCTTGTCACTGTTACACTGAAAGTGCACACGCATGACAAAATG 2244

TAGACGAGACGCCTCAAGGTACTGGATGCAAGCAGGATTTTTGCCCTTTAGTTTTCCAAGACACCTTTCTTTTATTATG 2323

CACTTGAGACAAGAGAATTAATAGAGCGTTAATTCAACAGGAAGACCGCCTCCAACCAAAGACCTGGAGCGCAGCATAA 2402

GGACTTGTGATTTGAGACGTTGTCCCCAGCCTGGTAGATCCCCCTTTCTCAGCATTTGGGATTTAGCAGTGCATAAAGC 2481

ATTAATATCTGTAAAAACACCTAGATGTTTGTTTGGCTTTTAATTTAAGGAAGCTGCAACCACAAAGCTTCCGCTCAGG 2560

GTTTTTTCTTCCTTCAAGTCTCCAAGGGCTCTTCAGCGTCACAAGCCAGCAACTCTCTTTGCATGAAAATTTCAAAGTT 2639

TAATTAATAATTAAAGGCAACAGCAAGCAGCAGCCTGTGAAGATTTTGCTCATCTTTTTTATGCCTTTTGATATTGAGT 2718

GACCTATCACTGTATGCATGTTACTTAGAAATTGAGGAGCACCACACCTTGGTTGTGGTTCAGCCTGGGAAAGAGACCT 2797

CCTTCCTTCTGTTTATAAATTAAAATCAGGAGGGGCGCCATCAGAAAGCATGGACAATATACATACTATAAATTTTTAG 2876

AAATATCACCATCGTGTCACGTCAACGATGCCAAATTATGTTAGTGTGAGCAGAAACCCGGTGGGGGAGGAAGGCGGCA 2955

GCAGCCGAAGGAAATAGCTCAGATAATCTAGTCACTTTCGATACTGTACTTCAGATGCGAAATGGATATTCGACTGGAA 3034

ACCTGACAAAGCGCGCCTGCTTTGATGTGAACTGTTATAGACAATGACCAGTGGCCGGGTCAGTGGGATGTCTCTCTGC 3113

GAGCACAAAGGCTTATCAAATGACACTAAAAATAAGTTCAACAACCATCACTTTGGAAGGGAGAAGGCGAACATTTCAT 3192

GTTTGGCGGGCATGTGAGTGCAGGAGATGGAAAGAGCCATTTAGAGCATCCTCATATAATTGCCCGCATTGTACTCTTC 3271

ATGGAAATTTCAAAGGACGGCAGTGATATTTTTCATTGGTGTCCACGTTTGTGGCACTGCTCCAAGAAGCCTTATGCAC 3350

ACATACAAATATACAAATGCACACACCTACACTCTCTAGCTTTAATCTTTTTGCTCAACCTTATTTATATCACTGACTG 3429

GCTGGATCCAAAGTCACACCTCCACATATTAGTGAATAAAAAATTTTTACCTGTGAAAAAAAAAAAAAAAAA 3501

FIG.2C

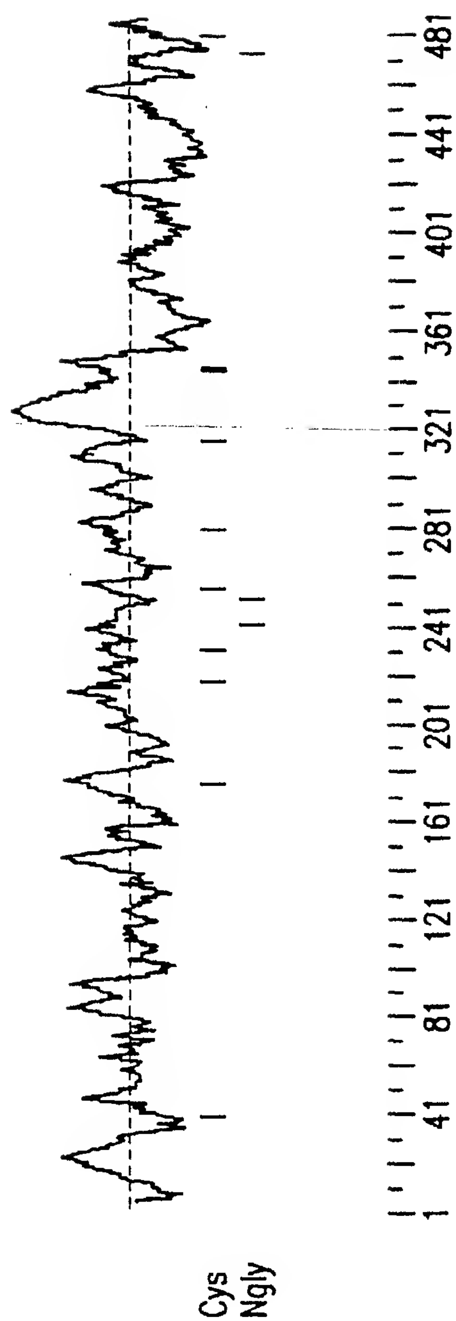


FIG.3

```
  1 MVAERSPARSPGSWLFPGLWLLVLSGPGGLLRAQEQPSCRRAFDLYFVLD  50
    ||| || ||||||||||||||| . ||| ||.|||||||::|||||||||
  1 MVAGRSRARSPGSWLFPGLWLLAVGGPGSLLQAQEQPSCKKAFDLYFVLD  50

 51 KSGSVANNWIEIYNFVQQLAERFVSPEMRLSFIVFSSQATIILPLTGDRG 100
    |||||||||||||||| || ||||||||||||||||||||||||||||| 
 51 KSGSVANNWIEIYNFVHQLTERFVSPEMRLSFIVFSSQATIILPLTGDRY 100

101 KISKGLEDLKRVSPVGETYIHEGLKLANEQIQKAGGLKTSSIIIALTDGK 150
    || ||||||| | ||||||||||||||||||| ||||| |||||||||||
101 KIGKGLEDLKAVKPVGETYIHEGLKLANEQIQNAGGLKASSIIIALTDGK 150

151 LDGLVPSYAEKEAKISRSLGASVYCVGVLDFEQAQLERIADSKEQVFPVK 200
    |||||||||| ||| ||||||||||||||||||||||||||||:||||||
151 LDGLVPSYAENEAKKSRSLGASVYCVGVLDFEQAQLERIADSKDQVFPVK 200

201 GGFQALKGIINSILAQSCTEILELQPSSVCVGEEFQIVLSGRGFMLGSRN 250
    |||||||||||||||||||||||| ||||||||.||:||.||    | . 
201 GGFQALKGIINSILAQSCTEILELSPSSVCVGEKFQVVLTGRAVTSISHD 250

251 GSVLCTYTVNETYTTSVKPVSVQLNSMLCPAPILNKAGETLDVSVSFNGG 300
    ||||||:| | ||| | ||||:| .|.|||||:||| ||||:||:|:| |
251 GSVLCTFTANSTYTKSEKPVSIQPSSILCPAPVLNKDGETLEVSISYNDG 300

301 KSVISGSLIVTATECSNGIAAIIVILVLLLLLGIGLMWWFWPLCCKVVIK 350
    || :| || :|||||.||||||: |||||||||  |||||||||||||||
301 KSAVSRSLTITATECTNGIAAIVAILVLLLLLGAALMWWFWPLCCKVVIK 350

351 DPPPPPPPAPKEEEEEPLPTKKWPTVDASYYGGRGVGGIKRMEVRWGDKG 400
    ||||||    .||||:||| ||||||||||||||||||||||||||||||
351 DPPPPPSAPMEEEEEDPLPNKKWPTVDASYYGGRGVGGIKRMEVRWGDKG 400

401 STEEGARLEKAKNAVVKIPEETEEPIRPRPPRPKPTHQPPQTKWYTPIKG 450
    |||||||||||||||| :||| | ||  |||||:|||| |||||||||||
401 STEEGARLEKAKNAVVMVPEE.EIPIPSRPPRPRPTHQAPQTKWYTPIKG 449

451 RLDALWALLRRQYDRVSLMRPQEGDEGRCINFSRVPSQ 488
    ||||||||: :||||||||||||||||||||||||| |
450 RLDALWALIMKQYDRVSLMRPQEGDEGRCINFSRVPHQ 487
```

FIG.4

```
      M   A   E   L   E   F   V   Q   I   I   I   I   V   V   V   M   M   V    18
GGGAG ATG GCG GAG CTG GAG TTT GTT CAG ATC ATC ATC ATC GTG GTG GTG ATG ATG GTG  59

M   V   V   V   I   T   C   L   L   S   H   Y   K   L   S   A   R   S   F   I    38
ATG GTG GTG GTG ATC ACG TGC CTG CTG AGC CAC TAC AAG CTG TCT GCA CGG TCC TTC ATC 119

S   R   H   S   Q   G   R   R   R   E   D   A   L   S   S   E   G   C   L   W    58
AGC CGG CAC AGC CAG GGG CGG AGG AGA GAA GAT GCC CTG TCC TCA GAA GGA TGC CTG TGG 179

P   S   E   S   T   V   S   G   N   G   I   P   E   P   Q   V   Y   A   P   P    78
CCC TCG GAG AGC ACA GTG TCA GGC AAC GGA ATC CCA GAG CCG CAG GTC TAC GCC CCG CCT 239

R   P   T   D   R   L   A   V   P   P   F   A   Q   R   E   R   F   H   R   F    98
CGG CCC ACC GAC CGC CTG GCC GTG CCG CCC TTC GCC CAG CGG GAG CGC TTC CAC CGC TTC 299

Q   P   T   Y   P   Y   L   Q   H   E   I   D   L   P   P   T   I   S   L   S   118
CAG CCC ACC TAT CCG TAC CTG CAG CAC GAG ATC GAC CTG CCG CCC ACC ATC TCG CTG TCA 359

D   G   E   E   P   P   P   Y   Q   G   P   C   T   L   Q   L   R   D   P   E   138
GAC GGG GAG GAG CCC CCA CCC TAC CAG GGC CCC TGC ACC CTC CAG CTT CGG GAC CCC GAG 419

Q   Q   L   E   L   N   R   E   S   V   R   A   P   P   N   R   T   I   F   D   158
CAG CAG CTG GAA CTG AAC CGG GAG TCG GTG CGC GCA CCC CCA AAC AGA ACC ATC TTC GAC 479

S   D   L   M   D   S   A   R   L   G   G   P   C   P   P   S   S   N   S   G   178
AGT GAC CTG ATG GAT AGT GCC AGG CTG GGC GGC CCC TGC CCC CCC AGC AGT AAC TCG GGC 539

I   S   A   T   C   Y   G   S   G   G   R   M   E   G   P   P   P   T   Y   S   198
ATC AGC GCC ACG TGC TAC GGC AGC GGC GGG CGC ATG GAG GGG CCG CCG CCC ACC TAC AGC 599

E   V   I   G   H   Y   P   G   S   S   F   Q   H   Q   Q   S   S   G   P   P   218
GAG GTC ATC GGC CAC TAC CCG GGG TCC TCC TTC CAG CAC CAG CAG AGC AGT GGG CCG CCC 659

S   L   L   E   G   T   R   L   H   H   T   H   I   A   P   L   E   S   A   A   238
TCC TTG CTG GAG GGG ACC CGG CTC CAC CAC ACA CAC ATC GCG CCC CTA GAG AGC GCA GCC 719

I   W   S   K   E   K   D   K   Q   K   G   H   P   L   *                       253
ATC TGG AGC AAA GAG AAG GAT AAA CAG AAA GGA CAC CCT CTC TAG                     764

GGTCCCCAGGGGGGCCGGGCTGGGGCTGCGTAGGTGAAAAGGCAGAACACTCCGCGCTTCTTAGAAGAGGAGTGAGAGG 843

AAGGCGGGGGGCGCAGCAACGCATCGTGTGGCCCTCCCCTCCCACCTCCCTGTGTATAAATATTTACATGTGATGTCTG 922

GTCTGAATGCACAAGCTAAGAGAGCTTGCAAAAAAAAAAAAAAAAAA                                  969
```

FIG.5

```
  L   S   H   Y   K   L   S   A   R   S   F   I   S   R   H   S   Q   A   R      19
G CTG AGC CAC TAC AAG CTG TCA GCC CGC TCC TTC ATC AGC CGA CAC AGC CAG GCC AGG     58

 R   R   D   D   G   L   S   S   E   G   C   L   W   P   S   E   S   T   V   S    39
AGG AGA GAC GAT GGA CTG TCC TCG GAA GGA TGC CTC TGG CCC TCA GAG AGT ACG GTG TCA  118

 G   G   M   P   E   P   Q   V   Y   A   P   P   R   P   T   D   R   L   A   V    59
GGT GGA ATG CCG GAG CCA CAG GTC TAT GCC CCG CCT CGG CCC ACT GAC CGA CTC GCT GTG  178

 P   P   F   I   Q   R   S   R   F   Q   P   T   Y   P   Y   L   Q   H   E   I    79
CCC CCC TTC ATC CAG CGG AGC CGA TTC CAA CCC ACC TAC CCC TAC CTG CAG CAC GAA ATT  238

 A   L   P   P   T   I   S   L   S   D   G   E   E   P   P   P   Y   Q   G   P    99
GCC CTG CCA CCC ACC ATC TCA CTG TCT GAT GGG GAG GAG CCC CCA CCC TAC CAG GGC CCC  298

 C   T   L   Q   L   R   D   P   E   Q   Q   L   E   L   N   R   E   S   V   R   119
TGC ACC CTC CAG CTA CGG GAC CCT GAG CAA CAG CTG GAG CTG AAC CGG GAA TCT GTG CGC  358

 A   P   P   N   R   T   I   F   D   S   D   L   I   D   S   T   M   L   G   G   139
GCA CCC CCT AAC CGG ACC ATC TTC GAC AGT GAC CTT ATA GAC AGC ACC ATG CTG GGG GGC  418

 P   C   P   P   S   S   N   S   G   I   S   A   T   C   Y   S   S   G   G   R   159
CCC TGT CCC CCC AGC AGT AAC TCG GGC ATC AGC GCC ACC TGC TAC AGC AGC GGT GGG CGC  478

 M   E   G   P   P   P   T   Y   S   E   V   I   G   H   Y   P   G   S   S   F   179
ATG GAG GGG CCG CCC CCC ACC TAC AGC GAG GTC ATT GGC CAC TAC CCT GGC TCC TCC TTC  538

 Q   H   Q   Q   S   N   G   P   S   S   L   L   E   G   T   R   L   H   H   S   199
CAG CAC CAG CAA AGT AAC GGG CCA TCC TCC CTG CTA GAG GGG ACC CGG CTC CAT CAC TCG  598

 H   I   A   P   L   E   N   K   E   K   E   K   Q   K   G   H   P   L   *       218
CAC ATT GCC CCA CTG GAG AAC AAG GAG AAG GAG AAA CAG AAA GGT CAC CCC CTC TAG      655
```

FIG.6A

GAGTGGGGCCGGGGCGCCTGTAGGCAAAACCGCAAAAAAAAAAAGAAAGGAAAAAAAAAACACTCCGCACTTCTTAAGA 734

GAGAAGAGAGAGGAAGTCAGGGGACACACAGGCTGAGTGGCCGTGTGGTAGTTCTCCTGTCTGTGTATAAATATTTACA 813

TGTTCTGTGTGGTCTGAATGCAGAAGCTCAAAAAGCTTGCAAAAAAGAAAAAAGAAAATGTTTCTTTGTTGAGCCGTGT 892

CTTGAAGGCAAAAGAGAAAATGCTCTACTAGTCTTTCCTTCATTGTTCTTAGTTGAGCTCTGTGCGTGAATGCTTAATT 971

TCTTTTGTTTATGACGGTTTTACTTAACTTTAAAGACATATTTGCACAAAACCTTTGTTTAAGATCTGCAATATTATAT 1050

ATATAAATATATATATAAAATAAGAGAACTCTGTATGTGGGAGAGCAGGAGTATTTTTGTATTAGAAGAGGCCTATTCA 1129

AAAAAAAAAAAGTTGTTTTCTGAACTAGAAGAGAAAAAAAATGGCAATTTTTTGAGTGCCAACTCCGAAAGTGTGTATT 1208

ACCTTGTAAAGAAAAAAAAAACCTACAAAGCAGGGGTTTAGAGCTATTTATATAAATGCTGCGATTTTGCACTATTTTT 1287

TAATATAATATATCAGTGCTTGTTTGATGGAAACTTCTAGAGAGTGTGTTGAGACTTGGAGAAATGTTGAGTTTAAGTC 1366

TGAACAGAGCAGCAGGGGTGTGTCCCTCCACCCCCTGGGTCTCAGAGGGCCCCCACTGTGCAGGAAAGCCAGGCTCCCC 1445

CAGGGCCACACAGCATGCCGCCTGCAAGCCAATGGATGCTTCTGTCCAGACAGGCTGTGAACAGTTACACATTCCCTGC 1524

CGAAGGGGACCCGTGATGTGTCGTCCCTCTGGGTTACAGCTGAGTCAGGACCGTCTCCCCCGCACCGCCCCCCAACAGA 1603

GGCATCCAATTCCAGATGCCTTTCCCAATATGAAAATAAACTGTTACCGAAGGAATCTTAAAAAAAAAAAAAAAAAAAA 1682

AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1713

FIG.6B

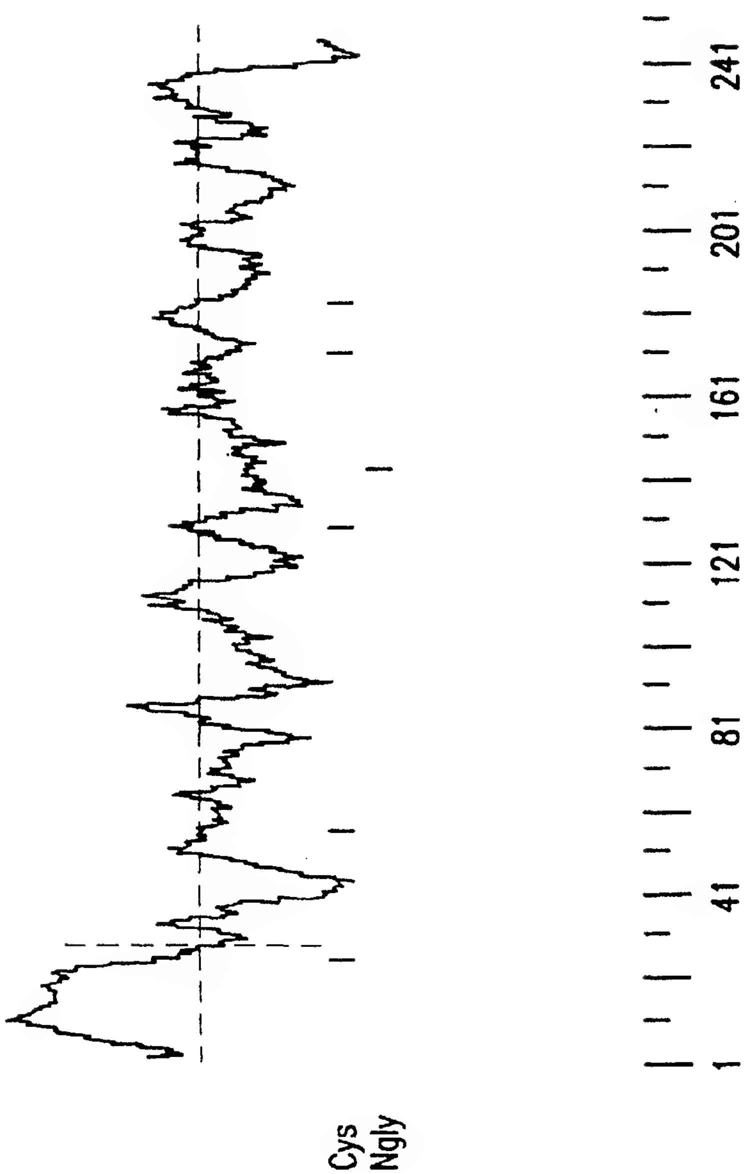


FIG.7

```
  1 MAELEFVQIIIIVVVMMVMVVVITCLLSHYKLSARSFISRHSQGRRREDA  50
                              ||||||||||||||||| |||:| 
  1 ..........................LSHYKLSARSFISRHSQARRRDDG  24

 51 LSSEGCLWPSESTVSGNGIPEPQVYAPPRPTDRLAVPPFAQRERFHRFQP 100
    |||||||||||||||| |.|||||||||||||||||||| ||    ||||
 25 LSSEGCLWPSESTVSG.GMPEPQVYAPPRPTDRLAVPPFIQRS...RFQP  70

101 TYPYLQHEIDLPPTISLSDGEEPPPYQGPCTLQLRDPEQQLELNRESVRA 150
    ||||||||| ||||||||||||||||||||||||||||||||||||||||
 71 TYPYLQHEIALPPTISLSDGEEPPPYQGPCTLQLRDPEQQLELNRESVRA 120

151 PPNRTIFDSDLMDSARLGGPCPPSSNSGISATCYGSGGRMEGPPPTYSEV 200
    |||||||||||.||  |||||||||||||||||| |||||||||||||||
121 PPNRTIFDSDLIDSTMLGGPCPPSSNSGISATCYSSGGRMEGPPPTYSEV 170

201 IGHYPGSSFQHQQSSGPPSLLEGTRLHHTHIAPLESAAIWSKEKDKQKGH 250
    ||||||||||||||.|| ||||||||||.||||||     .|||:|||||
171 IGHYPGSSFQHQQSNGPSSLLEGTRLHHSHIAPLE.....NKEKEKQKGH 215

251 PL 252
    ||
216 PL 217
```

FIG.8

GCGGTGACGGCGGCGCCTGGCGGCGGGGATTTGGCGCGGGCCGGGGGCCGGGGGCCGGGGCG 62

CGGGGGCGCGAGGCTGGATTCCTAGGGCCGCGGCGCTTCCCGGCATGCTCCGCTGCAGGCCCGCGCCCGCGCCCGGACT 141

TTGCCATCGGCGGGGCAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGTCTCTGCAGGTGTCGTGGAG 220

GAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCAGCAGCTTTAGCCCATGAGGAGGATGTGACCGGGACT 299

M E T V V I V A I G V L A T 14
GAGTCAGGAGCCCTCTGGAAGC ATG GAG ACT GTG GTG ATT GTT GCC ATA GGT GTG CTG GCC ACC 363

I F L A S F A A L V L V C R Q R Y C R P 34
ATC TTT CTG GCT TCG TTT GCA GCC TTG GTG CTG GTT TGC AGG CAG CGC TAC TGC CGG CCG 423

R D L L Q R Y D S K P I V D L I G A M E 54
CGA GAC CTG CTG CAG CGC TAT GAT TCT AAG CCC ATT GTG GAC CTC ATT GGT GCC ATG GAG 483

T Q S E P S E L E L D D V V I T N P H I 74
ACC CAG TCT GAG CCC TCT GAG TTA GAA CTG GAC GAT GTC GTT ATC ACC AAC CCC CAC ATT 543

E A I L E N E D W I E D A S G L M S H C 94
GAG GCC ATT CTG GAG AAT GAA GAC TGG ATC GAA GAT GCC TCG GGT CTC ATG TCC CAC TGC 603

I A I L K I C H T L T E K L V A M T M G 114
ATT GCC ATC TTG AAG ATT TGT CAC ACT CTG ACA GAG AAG CTT GTT GCC ATG ACA ATG GGC 663

S G A K M K T S A S V S D I I V V A K R 134
TCT GGG GCC AAG ATG AAG ACT TCA GCC AGT GTC AGC GAC ATC ATT GTG GTG GCC AAG CGG 723

I S P R V D D V V K S M Y P P L D P K L 154
ATC AGC CCC AGG GTG GAT GAT GTT GTG AAG TCG ATG TAC CCT CCG TTG GAC CCC AAA CTC 783

L D A R T T A L L L S V S H L V L V T R 174
CTG GAC GCA CGG ACG ACT GCC CTG CTC CTG TCT GTC AGT CAC CTG GTG CTG GTG ACA AGG 843

N A C H L T G G L D W I D Q S L S A A E 194
AAT GCC TGC CAT CTG ACG GGA GGC CTG GAC TGG ATT GAC CAG TCT CTG TCG GCT GCT GAG 903

E H L E V L R E A A L A S E P D K G L P 214
GAG CAT TTG GAA GTC CTT CGA GAA GCA GCC CTA GCT TCT GAG CCA GAT AAA GGC CTC CCA 963

FIG.9A

G P E G F L Q E Q S A I * 227
GGC CCT GAA GGC TTC CTG CAG GAG CAG TCT GCA ATT TAG 1002

TGCCTACAGGCCAGCAGCTAGCCATGAAGGCCCCTGCCGCCATCCCTGGATGGCTCAGCTTAGCCTTCTACTTTTTCCT 1081

ATAGAGTTAGTTGTTCTCCATGGCTGGAGAGTTCAGCTGTGTGTGCATAGTAAAGCAGGAGATCCCCGTCAGTTTATGC 1160

CTCTTTTGCAGTTGCAAACTGTGGCTGGTGAGTGGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCA 1239

AGAGGAGTATTGAAAACTGGTGGACTGTCAGCTTTATTTAGCTCACCTAGTGTTTTCAAGAAAATTGAGCCACCGTCTA 1318

AGAAATCAAGAGGTTTCACATTAAAATTAGAATTTCTGGCCTCTCTCGATTGGTCAGAATGTGTGGCAATTCTGATCTG 1397

CATTTTCAGAAGAGGACAATCAATTGAAACTAAGTAGGGGTTTCTTCTTTTGGCAAGACTTGTACTCTCTCACCTGGCC 1476

TGTTTCATTTATTTGTATTATCTGCCTGGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAGGTTTGGGTTTGAAG 1555

CTGAGGAACTACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTGGATAGTA 1634

AATTTATACTTATGTTTCCCTCAAAAAAAAAAAAAAAAAAAAAAAAAA 1682

FIG.9B

```
                         GCGGGGAGCCGGCGTACCCCGGAGCTGCAATTGAGCTTCCACCTGCTCAGAGGATGTGAGTG  62

                            M   E   T   V   V   I   V   A   I   G   V   L   A    13
GACTGAGGCTGGGAGCCCCTGGAAGC ATG GAG ACT GTG GTG ATC GTC GCC ATA GGT GTG CTG GCC  127

 T   I   F   L   A   S   F   A   A   L   V   V   V   C   R   Q   R   Y   C   R    33
ACC ATT TTC CTG GCC TCA TTT GCT GCT TTG GTG GTG GTC TGT AGG CAG CGT TAC TGC CGG  187

 P   R   D   L   L   Q   R   Y   D   S   K   P   I   V   D   L   I   G   A   M    53
CCT CGA GAC CTA CTA CAG CGT TAT GAT TCC AAG CCC ATC GTG GAC CTC ATT GGT GCT ATG  247

 E   T   Q   S   E   P   S   E   L   E   L   D   D   V   V   I   T   N   P   H    73
GAG ACG CAG TCT GAA CCC TCC GAG CTG GAG CTG GAT GAT GTC GTC ATC ACC AAC CCC CAC  307

 I   E   A   I   L   E   N   E   D   W   I   E   D   A   S   G   L   M   S   H    93
ATC GAG GCC ATC CTG GAG AAT GAG GAC TGG ATC GAA GAT GCC TCG GGC CTC ATG TCC CAC  367

 C   I   A   I   L   K   I   C   H   T   L   T   E   K   L   V   A   M   T   M   113
TGC ATC GCC ATC TTG AAG ATT TGT CAC ACT CTG ACA GAA AAA CTC GTT GCC ATG ACA ATG  427

 G   S   G   A   K   M   K   T   S   A   S   V   S   D   I   I   V   V   A   K   133
GGT TCT GGG GCC AAG ATG AAG ACG TCA GCA AGT GTC AGT GAC ATC ATT GTG GTG GCC AAA  487

 R   I   S   P   R   V   D   D   V   V   K   S   M   Y   P   P   L   D   P   K   153
CGG ATT AGC CCC AGA GTG GAC GAC GTC GTG AAG TCA ATG TAC CCT CCA CTG GAC CCC AAG  547

 L   L   D   A   R   T   T   A   L   L   L   S   V   S   H   L   V   L   V   T   173
CTC CTG GAT GCA CGG ACA ACC GCC CTG CTG CTG TCC GTT AGT CAC TTG GTG CTA GTG ACC  607

 R   N   A   C   H   L   T   G   G   L   D   W   I   D   Q   S   L   S   A   A   193
AGG AAC GCC TGC CAT CTA ACC GGG GGC CTG GAC TGG ATT GAC CAA TCA CTG TCT GCC GCT  667

 E   E   H   L   E   V   L   R   E   A   A   L   A   S   E   P   D   K   S   L   213
GAG GAG CAC CTG GAA GTC CTT CGA GAG GCA GCC CTG GCT TCT GAG CCA GAT AAA AGC CTC  727

 P   N   P   E   G   F   L   Q   E   Q   S   A   I   *                           227
CCC AAC CCT GAG GGC TTC CTG CAG GAA CAG TCG GCC ATT TAA                          769
```

FIG.10A

TCATCTCCGAGGCCCCATTTCCGCCCCTGGGCGAGCCTTCTACTTCCTGTAGATTTAGTTGTTCTCTAGAGCTCTGTCG 848

GCCAGCCCTGGGTGCAGGGTAAAGCCGAGAGCCTCACGCTGGACAGGCTCTGCTGCAATGGCAAACAGTGGCTGGAGAG 927

TGGCAGTGTAATCCCACAGTTAGGGGAGACGCTGTGTACCTCTACAGCAGAGCGCAGAAAGCTGCCGGGCTCGCTGCTA 1006

CATTTAGTTCATTTAATGTTTCCAAGAAAATCGAGTTGCCCTCTAAGAATTGAGAGACTTCATATCAAATTAGAATTTC 1085

CGGCTTCTGAAAATCAAGGCGTGGCAACATGGACAATCAGAACTAAGTGGCTAGGTCGAGATAGTCTCTTCGGGTGACC 1164

CTTGTGCTCCCCTTGCTAATTTGTTTGTGTTCAGTGCTTTGGTTCCTGAAGCATCAGAGCTCCCCACCCCCACCCCGCT 1243

TCCCATGTATCTTCCGCTTTCCCATTTGTGTTAGAAGCTGAGGAAATGCGAAGTCAATTGTTTCCTTTTTATCACTATG 1322

CCTGCAATTTCGCTTTACAACCACCAGGCGAATAGTAAACTTGTTCTTCTGTTTTCCTCAAAAAAAAAAAAAAAAAAAA 1401

AAAAAAAAAAAAAAAAAAAAAAAA 1425

FIG.10B

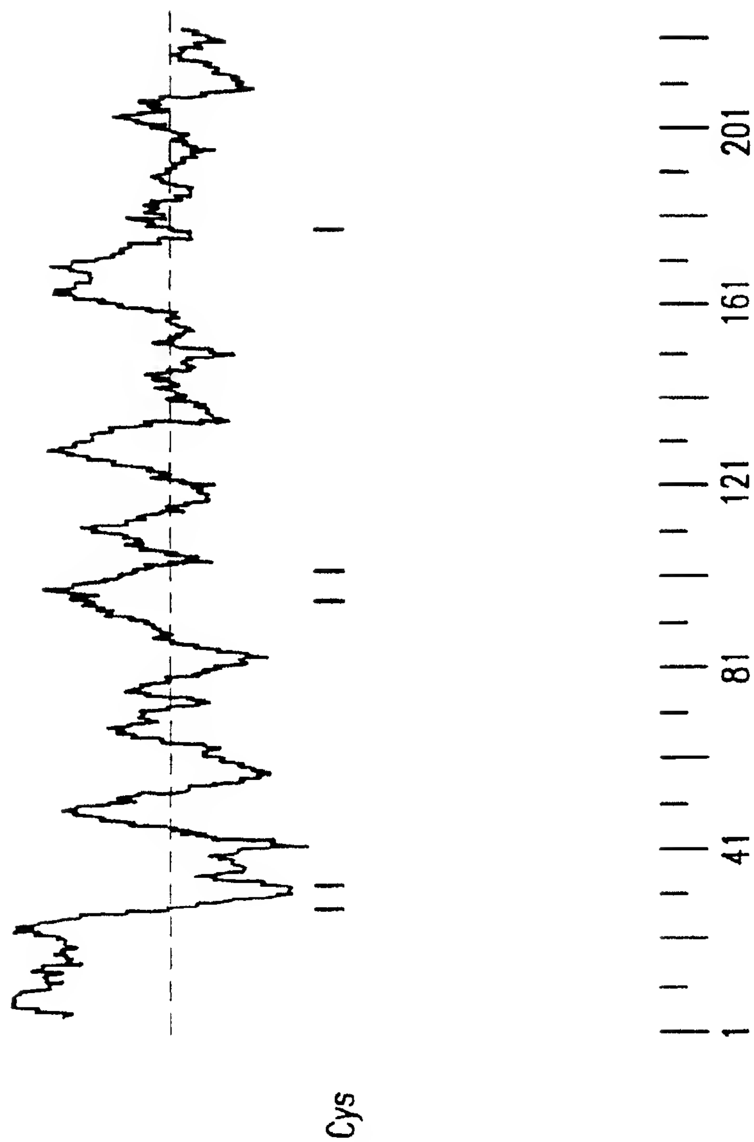


FIG. 11

```
  1 METVVIVAIGVLATIFLASFAALVLVCRQRYCRPRDLLQRYDSKPIVDLI  50
    ||||||||||||||||||||||||.|||||||||||||||||||||||||
  1 METVVIVAIGVLATIFLASFAALVVVCRQRYCRPRDLLQRYDSKPIVDLI  50

 51 GAMETQSEPSELELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKI 100
    ||||||||||||||||||||||||||||||||||||||||||||||||||
 51 GAMETQSEPSELELDDVVITNPHIEAILENEDWIEDASGLMSHCIAILKI 100

101 CHTLTEKLVAMTMGSGAKMKTSASVSDIIVVAKRISPRVDDVVKSMYPPL 150
    ||||||||||||||||||||||||||||||||||||||||||||||||||
101 CHTLTEKLVAMTMGSGAKMKTSASVSDIIVVAKRISPRVDDVVKSMYPPL 150

151 DPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWIDQSLSAAEEHLEVL 200
    ||||||||||||||||||||||||||||||||||||||||||||||||||
151 DPKLLDARTTALLLSVSHLVLVTRNACHLTGGLDWIDQSLSAAEEHLEVL 200

201 REAALASEPDKGLPGPEGFLQEQSAI 226
    ||||||||||| || |||||||||||
201 REAALASEPDKSLPNPEGFLQEQSAI 226
```

FIG.12

```
         10        20        30        40        50        60        70
METVVIVAIGVLATIFLASFAALVLVC-RQRYCRPRD-LLQRYDSKPIVDLIGAMETQSEPSELELDDVVITNPHIE
MDIVVVLALVVLAFVFILSLLILVITCQRRRLSRAKSKILVPFETPPNRYSRNAAAINNIDDIETMEQL---NGLLE
          10        20        30        40        50        60        70

      80        90        100       110       120       130       140       150
AILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSASVSDIIVVAKRISPRVDDVVKSMYP-PLD
ELLDRNQWLFEARGIMQHVVAVLTISKSVTTKLSVIQLPTSP----SPFHDAINMAMRNIYPRFDDLVESLAAQPID
        80        90        100       110            120       130       140

         160       170       180        190       200        210       220
PKLLDARTTALLLSVSHLVLVTRNACHLTGGL--DWIDQSLSAAEEHLEVLREAA-LASEPDKGLPGPEGFLQEQSA
MRLLEARALSLGTVCWSLYL----PYTLLDAKYKELMQKPLNEMNLHLVTIRTAAHLVSMAEKGA........
  150       160           170       180       190       200
```

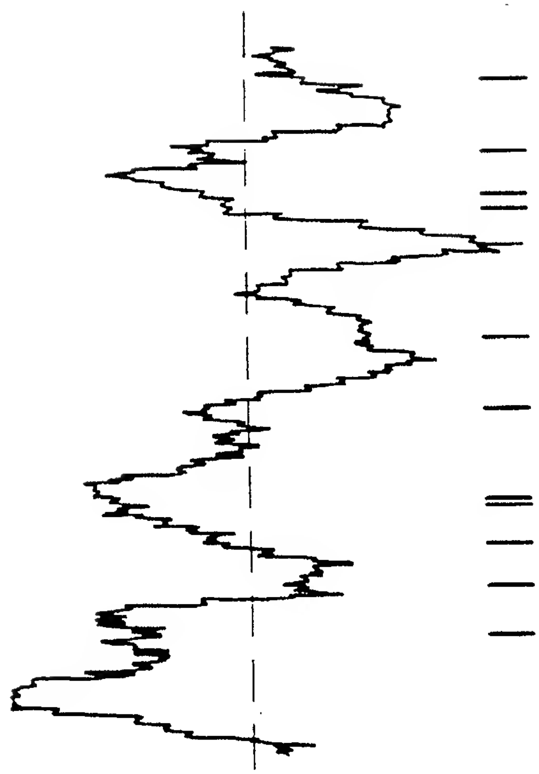
I

E

FIG.13

```
                                                 M   R   G     3
GCGAGAGGCATCTAAGCAGGCAGTGTTTTGCCTTCACCCCAAGTGACC ATG AGA GGT   57

 A   T   R   V   S   I   M   L   L   L   V   T   V   S   D   C   A   V   I   T    23
GCC ACG CGA GTC TCA ATC ATG CTC CTC CTA GTA ACT GTG TCT GAC TGT GCT GTG ATC ACA  117

 G   A   C   E   R   D   V   Q   C   G   A   G   T   C   C   A   I   S   L   W    43
GGG GCC TGT GAG CGG GAT GTC CAG TGT GGG GCA GGC ACC TGC TGT GCC ATC AGC CTG TGG  177

 L   R   G   L   R   M   C   T   P   L   G   R   E   G   E   E   C   H   P   G    63
CTT CGA GGG CTG CGG ATG TGC ACC CCG CTG GGG CGG GAA GGC GAG GAG TGC CAC CCC GGC  237

 S   H   K   V   P   F   F   R   K   R   K   H   H   T   C   P   C   L   P   N    83
AGC CAC AAG GTC CCC TTC TTC AGG AAA CGC AAG CAC CAC ACC TGT CCT TGC TTG CCC AAC  297

 L   L   C   S   R   F   P   D   G   R   Y   R   C   S   M   D   L   K   N   I   103
CTG CTG TGC TCC AGG TTC CCG GAC GGC AGG TAC CGC TGC TCC ATG GAC TTG AAG AAC ATC  357

 N   F   *                                                                       106
AAT TTT TAG                                                                      366

GCGCTTGCCTGGTCTCAGGATACCCACCATCCTTTTCCTGAGCACAGCCTGGATTTTTATTTCTGCCATGAAACCCAGC  445

TCCCATGACTCTCCCAGTCCCTACACTGACTACCCTGATCTCTCTTGTCTAGTACGCACATATGCACACAGGCAGACAT  524

ACCTCCCATCATGACATGGTCCCCAGGCTGGCCTGAGGATGTCACAGCTTGAGGCTGTGGTGTGAAAGGTGGCCAGCCT  603

GGTTCTCTTCCCTGCTCAGGCTGCCAGAGAGGTGGTAAATGGCAGAAAGGACATTCCCCCTCCCCTCCCCAGGTGACCT  682

GCTCTCTTTCCTGGGCCCTGCCCCTCTCCCCACATGTATCCCTCGGTCTGAATTAGACATTCCTGGGCACAGGCTCTTG  761

GGTGCATTGCTCAGAGTCCCAGGTCCTGGCCTGACCCTCAGGCCCTTCACGTGAGGTCTGTGAGGACCAATTTGTGGGT  840

AGTTCATCTTCCCTCGATTGGTTAACTCCTTAGTTTCAGACCACAGACTCAAGATTGGCTCTTCCCAGAGGGCAGCAGA  919

CAGTCACCCCAAGGCAGGTGTAGGGAGCCCAGGGAGGCCAATCAGCCCCCTGAAGACTCTGGTCCCAGTCAGCCTGTGG  998

CTTGTGGCCTGTGACCTGTGACCTTCTGCCAGAATTGTCATGCCTCTGAGGCCCCCTCTTACCACACTTTACCAGTTAA  1077

CCACTGAAGCCCCCAATTCCCACAGCTTTTCCATTAAAATGCAAATGGTGGTGGTTCAATCTAATCTGATATTGACATA  1156

TTAGAAGGCAATTAGGGTGTTTCCTTAAACAACTCCTTTCCAAGGATCAGCCCTGAGAGCAGGTTGGTGACTTTGAGGA  1235

GGGCAGTCCTCTGTCCAGATTGGGGTGGGAGCAAGGGACAGGGAGCAGGGCAGGGGCTGAAAGGGGCACTGATTCAGAC  1314

CAGGGAGGCAACTACACACCAACCTGCTGGCTTTAGAATAAAAGCACCAACTGAAAAAAAAAAAAAAAAAAAAAAAAAA  1393

AAAAAAAAAAAAAAAAAAAAAAAAAAAAA                                                    1422
```

FIG.14



Cys

FIG. 15

```
         10        20        30        40        50        60        70
         |         |         |         |         |         |         |
MRGATRVSIMLLLVTVSDCAVITGACERDVQCGAGTCCAISLWLRGLRMCTPLGREGEECHPGSHKVPFFRKRK-HH
                   ••••••••••|••• •• ••| ••||||•|•••|•  ••|•••|•••|••  |•• ••
-------------------AVITGACERDLQCGKGTSCAVCLWIKSVRVCTPVGTSGEDCHPASHKIPFSGQRKMHH
                             |         |         |         |         |
                             10        20        30        40        50

    80        90        100
    |         |         |
TCPCLPNLLCSRFPDGRYRCSMDLKNINF
•••• ••• • | |  |||•      ||
TCPCAPNLACVQTSPKKFKC------LSK
 |         |               |
 60        70              80
```

FIG.16

CGCTGGCATTTTCTCCTGGACAAGGAGAGAGTGCGGCTGCTGAGAGCCGAGCCCAGCAATCC 62

CGATCCTCTGAGTCGTGAAGAAGGGAGGCAGCGAGGGGGTTGGGGTTGGGGCCTGAGGCAAGCCCCCAGGCTCCGCTCT 141

M A Q K M D C G A G L L G F Q 15
TGCCAGAGGGACAGGAGCC ATG GCT CAG AAA ATG GAC TGT GGT GCG GGC CTC CTC GGC TTC CAG 205

A E A S V E D S A L L M Q T L M E A I Q 35
GCT GAG GCC TCC GTA GAA GAC AGC GCC TTG CTT ATG CAG ACC TTG ATG GAG GCC ATC CAG 265

I S E A P P T N Q A T A A A S P Q S S Q 55
ATC TCA GAG GCT CCA CCT ACT AAC CAG GCC ACC GCA GCT GCT AGT CCC CAG AGT TCA CAG 325

P P T A N E M A D I Q V S A A A A R P K 75
CCC CCA ACT GCC AAT GAG ATG GCT GAC ATT CAG GTT TCA GCA GCT GCC GCT AGG CCT AAG 385

S A F K V Q N A T T K G P N G V Y D F S 95
TCA GCC TTT AAA GTC CAG AAT GCC ACC ACA AAA GGC CCA AAT GGT GTC TAT GAT TTC TCT 445

Q A H N A K D V P N T Q P K A A F K S Q 115
CAG GCT CAT AAT GCC AAG GAT GTG CCC AAC ACG CAG CCC AAG GCA GCC TTT AAG TCC CAA 505

N A T P K G P N A A Y D F S Q A A T T G 135
AAT GCT ACC CCA AAG GGT CCA AAT GCT GCC TAT GAT TTT TCC CAG GCA GCA ACC ACT GGT 565

E L A A N K S E M A F K A Q N A T T K V 155
GAG TTA GCT GCT AAC AAG TCT GAG ATG GCC TTC AAG GCC CAG AAT GCC ACT ACT AAA GTG 625

G P N A T Y N F S Q S L N A N D L A N S 175
GGC CCA AAT GCC ACC TAC AAT TTC TCT CAG TCT CTC AAT GCC AAT GAC CTG GCC AAC AGC 685

R P K T P F K A W N D T T K A P T A D T 195
AGG CCT AAG ACC CCT TTC AAG GCT TGG AAT GAT ACC ACT AAG GCC CCA ACA GCT GAT ACC 745

Q T Q N V N Q A K M A T S Q A D I E T D 215
CAG ACC CAG AAT GTA AAT CAG GCC AAA ATG GCC ACT TCC CAG GCT GAC ATA GAG ACC GAC 805

P G I S E P D G A T A Q T S A D G S Q A 235
CCA GGT ATC TCT GAA CCT GAC GGT GCA ACT GCA CAG ACA TCA GCA GAT GGT TCC CAG GCT 865

Q N L E S R T I I R G K R T R K I N N L 255
CAG AAT CTG GAG TCC CGG ACA ATA ATT CGG GGC AAG AGG ACC CGC AAG ATT AAT AAC TTG 925

N V E E N S S G D Q R R A P L A A G T W 275
AAT GTT GAA GAG AAC AGC AGT GGG GAT CAG AGG CGG GCC CCA CTG GCT GCA GGG ACC TGG 985

FIG.17A

R S A P V P V T T Q N P P G A P P N V L 295
AGG TCT GCA CCA GTT CCA GTG ACC ACT CAG AAC CCA CCT GGC GCA CCC CCC AAT GTG CTC 1045

W Q T P L A W Q N P S G W Q N Q T A R Q 315
TGG CAG ACG CCA TTG GCT TGG CAG AAC CCC TCA GGC TGG CAA AAC CAG ACA GCC AGG CAG 1105

T P P A R Q S P P A R Q T P P A W Q N P 335
ACC CCA CCA GCA CGT CAG AGC CCT CCA GCT AGG CAG ACC CCA CCA GCC TGG CAG AAC CCA 1165

V A W Q N P V I W P N P V I W Q N P V I 355
GTC GCT TGG CAG AAC CCA GTG ATT TGG CCA AAC CCA GTA ATC TGG CAG AAC CCA GTG ATC 1225

W P N P I V W P G P V V W P N P L A W Q 375
TGG CCA AAC CCC ATT GTC TGG CCC GGC CCT GTT GTC TGG CCG AAT CCA CTG GCC TGG CAG 1285

N P P G W Q T P P G W Q T P P G W Q G P 395
AAT CCA CCT GGA TGG CAG ACT CCA CCT GGA TGG CAG ACC CCA CCG GGC TGG CAG GGT CCT 1345

P D W Q G P P D W P L P P D W P L P P D 415
CCA GAC TGG CAA GGT CCT CCT GAC TGG CCG CTA CCA CCC GAC TGG CCA CTG CCA CCT GAT 1405

W P L P T D W P L P P D W I P A D W P I 435
TGG CCA CTT CCC ACT GAC TGG CCA CTA CCA CCT GAC TGG ATC CCC GCT GAT TGG CCA ATT 1465

P P D W Q N L R P S P N L R P S P N S R 455
CCA CCT GAC TGG CAG AAC CTG CGC CCC TCG CCT AAC CTG CGC CCT TCT CCC AAC TCG CGT 1525

A S Q N P G A A Q P R D V A L L Q E R A 475
GCC TCA CAG AAC CCA GGT GCT GCA CAG CCC CGA GAT GTG GCC CTT CTT CAG GAA AGA GCA 1585

N K L V K Y L M L K D Y T K V P I K R S 495
AAT AAG TTG GTC AAG TAC TTG ATG CTT AAG GAC TAC ACA AAG GTG CCC ATC AAG CGC TCA 1645

E M L R D I I R E Y T D V Y P E I I E R 515
GAA ATG CTG AGA GAT ATC ATC CGT GAA TAC ACT GAT GTT TAT CCA GAA ATC ATT GAA CGT 1705

A C F V L E K K F G I Q L K E I D K E E 535
GCA TGC TTT GTC CTA GAG AAG AAA TTT GGG ATT CAA CTG AAA GAA ATT GAC AAA GAA GAA 1765

H L Y I L I S T P E S L A G I L G T T K 555
CAC CTG TAT ATT CTC ATC AGT ACC CCC GAG TCC CTG GCT GGC ATA CTG GGA ACG ACC AAA 1825

D T P K L G L L L V I L G V I F M N G N 575
GAC ACA CCC AAG CTC GGT CTC CTC TTG GTG ATT CTG GGT GTC ATC TTC ATG AAT GGC AAC 1885

FIG.17B

```
R   A   S   E   A   V   L   W   E   A   L   R   K   M   G   L   R   P   G   V    595
CGT GCC AGT GAG GCT GTC CTC TGG GAG GCA CTA CGC AAG ATG GGA CTG CGT CCT GGG GTG  1945

R   H   P   L   L   G   D   L   R   K   L   L   T   Y   E   F   V   K   Q   K    615
AGA CAT CCC CTC CTT GGA GAT CTA AGG AAA CTT CTC ACC TAT GAG TTT GTA AAG CAG AAA  2005

Y   L   D   Y   R   R   V   P   N   S   N   P   P   E   Y   E   F   L   W   G    635
TAC CTG GAC TAC AGA CGA GTG CCC AAC AGC AAC CCC CCG GAG TAT GAG TTC CTC TGG GGC  2065

L   R   S   Y   H   E   T   S   K   M   K   V   L   R   F   I   A   E   V   Q    655
CTC CGT TCC TAC CAT GAG ACT AGC AAG ATG AAA GTG CTG AGA TTC ATT GCA GAG GTT CAG  2125

K   R   D   P   R   D   W   T   A   Q   F   M   E   A   A   D   E   A   L   D    675
AAA AGA GAC CCT CGT GAC TGG ACT GCA CAG TTC ATG GAG GCT GCA GAT GAG GCC TTG GAT  2185

A   L   D   A   A   A   A   E   A   E   A   R   A   E   A   R   T   R   M   G    695
GCT CTG GAT GCT GCT GCA GCT GAG GCC GAA GCC CGG GCT GAA GCA AGA ACC CGC ATG GGA  2245

I   G   D   E   A   V   S   G   P   W   S   W   D   D   I   E   F   E   L   L    715
ATT GGA GAT GAG GCT GTG TCT GGG CCC TGG AGC TGG GAT GAC ATT GAG TTT GAG CTG CTG  2305

T   W   D   E   E   G   D   F   G   D   P   W   S   R   I   P   F   T   F   W    735
ACC TGG GAT GAG GAA GGA GAT TTT GGA GAT CCC TGG TCC AGA ATT CCA TTT ACC TTC TGG  2365

A   R   Y   H   Q   N   A   R   S   R   F   P   Q   T   F   A   G   P   I   I    755
GCC AGA TAC CAC CAG AAT GCC CGC TCC AGA TTC CCT CAG ACC TTT GCC GGT CCC ATT ATT  2425

G   P   G   G   T   A   S   A   N   F   A   A   N   F   G   A   I   G   F   F    775
GGT CCT GGT GGT ACA GCC AGT GCC AAC TTC GCT GCC AAC TTT GGT GCC ATT GGT TTC TTC  2485

W   V   E   *                                                                    779
TGG GTT GAG TGA                                                                  2497

GATGTTGGGTAGGTACATCACTTTGGATTGGGCAGTTAGGGTCTCTGGGGGATATAGGGTCCATGGGGTTGTATTTATG  2576

TGCATGAGCTAGAAGTATTAGGAAACTCATTGTAGGGAGATGAGGAAAGTATGGGGCAGCACTCTTTGGTATGTATATT  2655

CCTTATTATTTGATATATATCATTGATTTTTTTACTTCTTCTTTCTTATGTTCACAGATATTGCTATCAATCGCAGTAG  2734

TCTTTCCCCTGTGTGAGGCTGAAGCCTCAGATTCCTTCTAAACACAGCTATCTAGAGAGCCACATCCTGTTGACTGAAA  2813

GTGGCATGCAAGATAAATTTATTTGCTGTTCCTTGTCTACTGCTTTTTTTCCCCTTGTGTGCTGTCAAGTTTTGGTATC  2892

AGAAATAAACATTGAAATTGCAAAGTGAAAAAA                                               2925
```

FIG.17C

```
        10        20        30        40        50        60        70
MAQKMDCGAGLLGFQAEASVEDSALLMQTLMEAIQISEAPPTNQATAAASPQSSQPPTANEMADIQVSAAAARPKSA
MSDTSESGAGLTRFQAEASEKDSSSMMQTLLTVTQNVEVPET--------PKASKALEVSE--DVKVSKASGVSKAT
        10        20        30        40                50         60

80        90       100       110       120       130       140       150
FKVQNATTKGPNGVYDFSQAHNAKDVPNTQPKAAFKSQNATPKGPNAAYDFSQAATTGELAANKSEMAFKAQNATTK
---------------EVSKTSEAREAPATQASST----------------TQLTDTQVLAAENKSLA--ADTKKQN
                 70        80                          90        100        110

    160       170       180       190       200       210       220       230
VGPNATYNFSQSLNANDLANSRPKTPFKAWNDTTKAPTADTQTQNVNQAKMATSQADTETDPGISEPDGATAQTSAD
ADPQAV---------------------------TMPATETKKVSHVADTKVNTKAQETEAAPSQAPADEPEPESAAA
                                     120       130       140       150       160
```

FIG.18A

```
          240       250       260       270       280       290       300
            |         |         |         |         |         |         |
GSQAQNLESRTIIRGKRTRKINNLNVEENSSGDQRRAPLAAGTWRSAPVPVTTQNPPGAPPNVLWQTPLAWQNPSGW
 •• |• ||•  |||•||••|||•| ••||•|•• |•| ||• |• |   |||              |•|| | •
QSQ-ENQDTRPKVKAKKARKVKHLDGEEDGSSDQSQASGTTGGRRVSKALMAS--------------MARRASRG-
              |         |         |         |         |                   |
             170       180       190       200       210                 220

 310       320       330       340       350       360       370       380
   |         |         |         |         |         |         |         |
QNQTARQTPPARQSPPARQTPPAWQNPVAWQNPVIWPNPVIWQNPVIWPNPIVWPGPVVWPNPLAWQNPPGWQTPPG
                         •|•     |•|     |     •|   |           •| ||• |
-------------------------PIA-----FWARRASRTRLAAWARRAL----------LSLRSPKA------
                                           |         |                   |
                                          230       240                 250

   390       400       410       420       430       440       450       460
     |         |         |         |         |         |         |         |
WQTPPGWQGPPDWQGPPDWPLPPDWPLPPDWPLPTDWPLPPDWIPADWPIPPDWQNLRPSPNLRPSPNSRASQNPGA
                                                        • |     • ||| ||••|• •
--------------------------------------------------------RRGKARRRAAKLQSSQEPEA
                                                                 |         |
                                                                260       270
```

FIG.18B

```
470       480       490       500       510       520       530
AQPRDVALLQERANKLVKYLMLKDYTKVPIKRSEMLRDIIREYTDVYPEIIERACFVLEKKFGIQLKEIDKEEHLYI
PPPRDVALLQGRANDLVKYLLAKDQTKIPIKRSDMLKDIIKEYTDVYPEIIERAGYSLEKVFGIQLKEIDKNDHLYI
  280       290       300       310       320       330       340

540       550       560       570       580       590       600       610
LISTPESL-AGILGTTKDTPKLGLLLVILGVIFMNGNRASEAVLWEALRKMGLRPGVRHPLLGDLRKLLTYEFVKQK
LLSTLEPTDAGILGTTKDSPKLGLLMVLLSIIFMNGNRSSEAVIWEVLRKLGLRPGIHHSLFGDVKKLITDEFVKQK
  350       360       370       380       390       400       410       420

  620       630       640       650       660       670       680       690
YLDYRRVPNSNPPEYEFLWGLRSYHETSKMKVLRFIAEVQKRDPRDWTAQFMEAADEALDALDAAAAEAEARAEART
YLDYARVPNSNPPEYEFFWGLRSYYETSKMKVLKFACKVQKKDPKEWAAQYREAMEADLKAAAEAAAEAKARAEIRA
    430       440       450       460       470       480       490       500
```

FIG.18C

```
             700       710                           720       730       740
               |         |                             |         |         |
RMGIG--DEAVSGPWSWDDIEFELLTWD--------------EEGDFGDPWSRIPFTFWARYHQNARSRFPQ-----
•••••    •  |•• |••| ||   •              |•|| |   |  |    •   || | •|
RMGIGLGSENAAGPCNWDEADIG--PWAKARIQAGAEAKAKAQESGSASTGASTSTNNSASASASTSGGFSAGASLT
             |         |           |         |         |         |         |
            510       520         530       540       550       560       570

 750       760       770
   |         |         |
-TFAGPIIGPGGTASANFAANFGAIGFFWVE
 •||  |||  • •|•| |||  •• ••
ATLTFGLFAGLGGAGASTSGSSGACGFSY-K
   |         |         |
  580       590       600
```

FIG.18D

```
Mbkn    MRGATRVSIMLLLV-------  14
mBv8-3  MGDPRCAPLLLLLLLPLLFT   20
fBv8    M---KCFAQIVVLLLVIAF-   16
VPRA    --------------------    0

Mbkn    -TVSDCAVITGACERDVQCG   33
mBv8-3  PPAGDAAVITGACDKDSQCG   40
fBv8    ---SHGAVITGACDKDVQCG   33
VPRA    ------AVITGACERDLQCG   14

Mbkn    AGTCCAISLWLRGLRMCTPL   53
mBv8-3  GGMCCAVSIWVKSIRICTPM   60
fBv8    SGTCCAASAWSRNIRFCIPL   53
VPRA    KGTSCAVCLWIKSVRVCTPV   34

Mbkn    GREGEECHPGSHKVPFF-RK   72
mBv8-3  GQVGDSCHPLTRKVPFWG-R   79
fBv8    GNSGEDCHPASHKVPYDG-K   72
VPRA    GTSGEDCHPASHKIPFSGQR   54

Mbkn    RKHHTCPCLPNLLCSRFPDG   92
mBv8-3  RMHHTCPCLPGLACLRTSFN   99
fBv8    RLSSLCPCKSGLTCSKSG-E   91
VPRA    KMHHTCPCAPNLACVQTSPK   74

Mbkn    RYRCSMDLKNINF         105
mBv8-3  RFIC-LARK             107
fBv8    KFKCS                  96
VPRA    KFKC-LS-K              81
```

FIG.19A

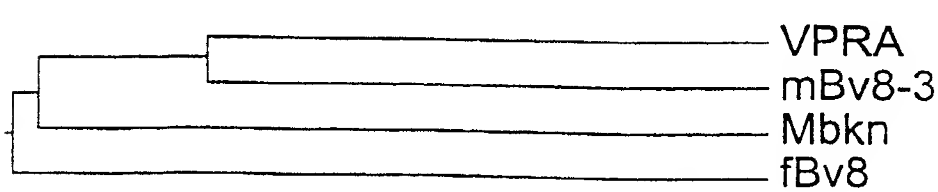


FIG.19B

FIG.19C

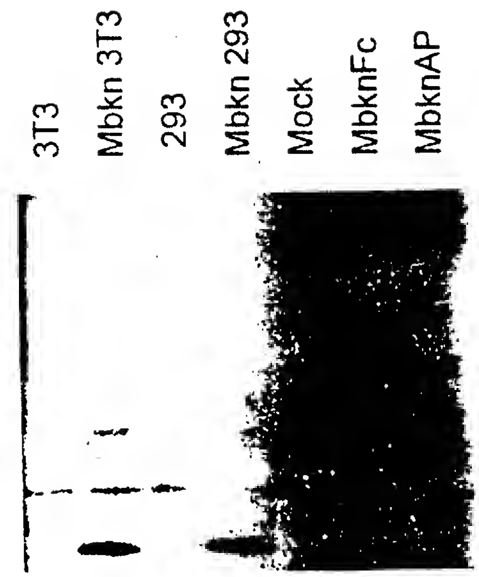


FIG.19D

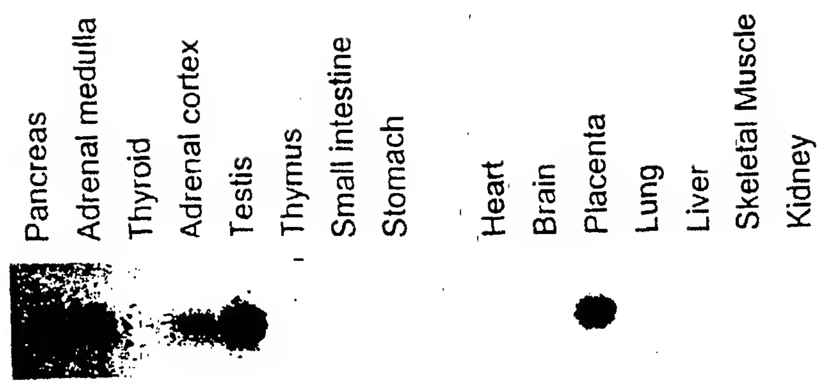


FIG.20A

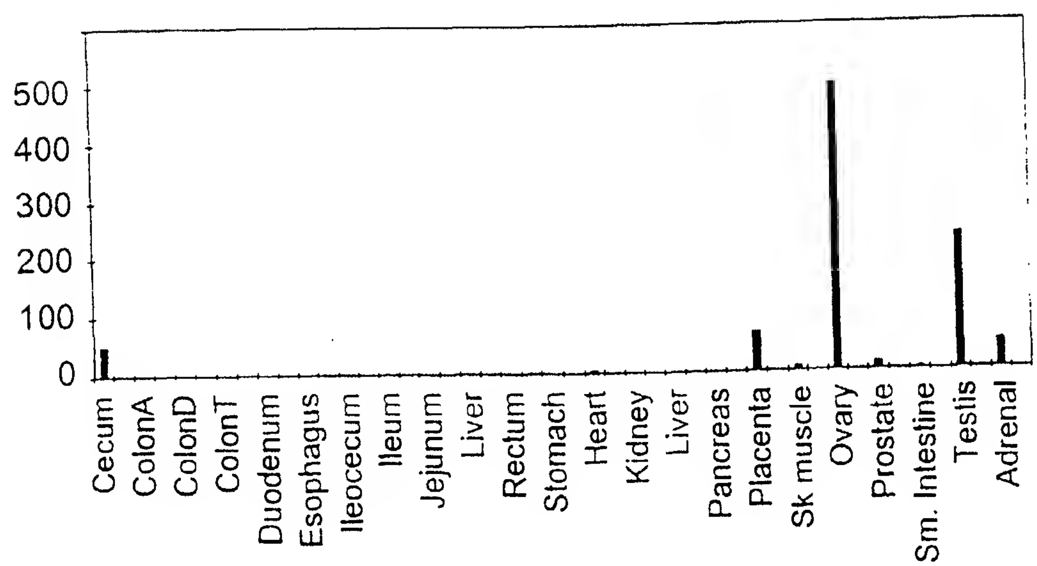


FIG.20B

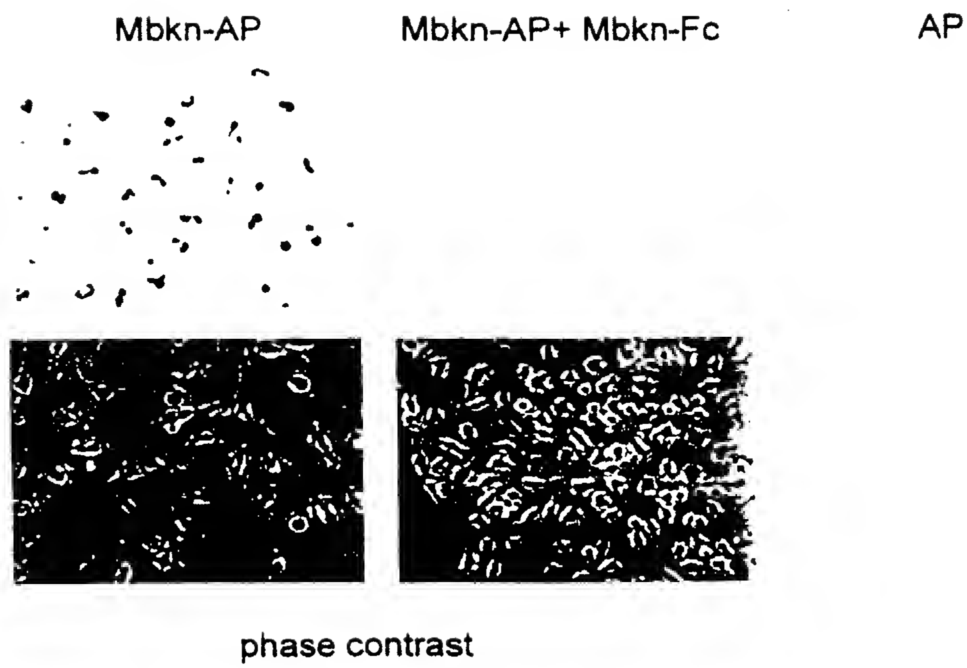


FIG.21

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05226

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) :Please See Extra Sheet.
US CL :Please See Extra Sheet.
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 536/23.1; 435/4, 7.1, 69.1, 320.1, 455, 325, 243; 530/350, 387.1

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

EAST: USPatfull, EPO,JPO, Derwent; STN: Medline, BIOSIS, Scisearch, Caplus; GenEMBL, N_Genseq_36, EST, PIR, SwissProt, SPTREMBL.

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X,P	Database EMBL, Accession No. HSM00861.30, KOEHRER et al., Homo sapiens cDNA from clone DKFZp586P1622. 30 August 1999, see entire document.	1-4

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
12 JULY 2000	11 AUG 2000

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231
Facsimile No. (703) 305-3230

Authorized officer
Nancy Ogihara
Cynthia Lawrence for
Telephone No. (703) 308-0196

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US00/05226

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This international report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

Please See Extra Sheet.

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☒ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:
1-18

Remark on Protest
☐ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

A. CLASSIFICATION OF SUBJECT MATTER:
IPC (7):

C07H 19/00, 21/00, 21/02, 21/04; C12Q 1/00; G01N 33.53; C12P 21/06, 15/09; C12N 15/00, 15/09, 15/63, 15/70, 15/74, 5/00, 5/02, 1/00, 15/63, 15/85, 15/87; C07K 1/00, 14/00, 17/00.

A. CLASSIFICATION OF SUBJECT MATTER:
US CL :

536/23.1; 435/4, 7.1, 69.1, 320.1, 455, 325, 243; 530/350, 387.1

BOX II. OBSERVATIONS WHERE UNITY OF INVENTION WAS LACKING
This ISA found multiple inventions as follows:

This application contains the following inventions or groups of inventions which are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for all inventions to be searched, the appropriate additional search fees must be paid.

Group I (claims 1-18), directed to an isolated nucleic acid and the corresponding polypeptide, methods of making the polynucleotide and the polypeptide, and methods of detecting the polynucleotide and polypeptide.
Group II (claims 19-20), directed to an a method of identifying a polypeptide.
Group III (claim 21), directed a method of modulating activity of a polypeptide.
Group IV (claim 22), directed to a method of identiyfing a compound that modulates activity of a peptide.

This application contains claims directed to more than one species of the generic invention. These species are deemed to lack Unity of Invention because they are not so linked as to form a single inventive concept under PCT Rule 13.1. In order for more than one species to be searched, the appropriate additional search fees must be paid. The species are as follows:

Species 1: human TANGO216 polypeptides (SEQ ID NO: 2) and polynucleotides.
Species 2: human TANGO261 polypeptides (SEQ ID NO: 5) and polynucleotides.
Species 3: human TANGO262 polypeptides (SEQ ID NO: 8) and polynucleotides.
Species 4: human TANGO266 polypeptides (SEQ ID NO: 11) and polynucleotides.
Species 5: human TANGO267 polypeptides (SEQ ID NO: 14) and polynucleotides.
Species 6: mouse TANGO216 polypeptides (SEQ ID NO: 17) and polynucleotides.
Species 7: mouse TANGO261 polypeptides (SEQ ID NO: 20) and polynucleotides.
Species 8: mouse TANGO262 polypeptides (SEQ ID NO: 23) and polynucleotides.

The claims are deemed to correspond to the species listed above in the following manner: the claims are directed toward isolated polynucleotides and polypeptides encoding TANGO polypeptides, methods of isolation, and methods of use.

The following claims are generic: 1-22

The inventions listed as Groups I-IV do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, they lack the same or corresponding special technical features for the following reasons: The special technical feature of Group I (claims 1-18) comprises a polynucleotide and polypeptide encoding a secreted TANGO protein, methods of producing, and methods of detecting the presence of the polynucleotide and the polypeptide. The methods of making and methods of detecting are not shared by Groups II-IV. Furthermore, Groups II-IV comprise binding compounds that modulate activity of the peptide which are neither required or encompassed by the claims of Invention I. Therefore, there is no shared concept or linking feature between the claims of Group I and Groups II-IV.

The species listed above do not relate to a single inventive concept under PCT Rule 13.1 because, under PCT Rule 13.2, the species lack the same or corresponding special technical features for the following reasons: The different species do not share a special technical feature since each encode differing proteins with differing sequences, with differing structures and functions, and differing biological roles.

A proper response to this request consists of an elected group and an elected species. In the absence of a response

INTERNATIONAL SEARCH REPORT

In ational application No.
PCT/US00/05226

Group I will be searched to the extent that it reads on Species 1. If more than one group and/or species is elected applicant should indicate in what manner they are to be combined.